



Recherche de la performance en simulation thermique dynamique : application à la réhabilitation des bâtiments

Mickael Rabouille

► To cite this version:

Mickael Rabouille. Recherche de la performance en simulation thermique dynamique : application à la réhabilitation des bâtiments. Génie civil. Université de Grenoble, 2014. Français. NNT : 2014GRENA024 . tel-01142344

HAL Id: tel-01142344

<https://theses.hal.science/tel-01142344>

Submitted on 15 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Génie Civil et Sciences de l'Habitat**

Arrêté ministériel : 7 août 2006

Présentée par

Mickael RABOUILLE

Thèse dirigée par **Etienne WURTZ**

codirigée par **Pascal PERROTIN**

préparée au sein du **LOCIE UMR 5271 CNRS-UDS**
dans l'**École Doctorale SISEO**

Recherche de la performance en simulation thermique dynamique : Application à la réhabilitation des bâtiments

Thèse soutenue publiquement le **07 novembre 2014**,
devant le jury composé de :

M, Monika, WOLOSZYN

Professeur, Université de Savoie, Présidente

M, Stéphane, LASSUE

Professeur, Université d'Artois, Rapporteur

M, Bruno, PEUPORTIER

Maître de recherche, MINES ParisTech, Rapporteur

M, Laurent, MORA

Maître de conférences, Université de Bordeaux1, Examineur

M, Etienne, WURTZ

Directeur de Recherche, CEA-INES, Directeur de thèse

M, Pascal, PERROTIN

Maître de conférences, Université de Savoie, Co-Directeur de thèse



REMERCIEMENTS

Je souhaite tout d’abord remercier mes directeurs de thèse Etienne WURTZ et Pascal PERROTIN pour m’avoir offert l’opportunité de réaliser ces travaux et m’avoir accompagné durant ces années. Merci à mon jury de thèse de m’avoir fait l’honneur d’évaluer mes travaux et particulièrement à mes rapporteurs Bruno PEUPORTIER et Stéphane LASSUE. Merci à Laurent MORA et Monika WOLOSZYN ma présidente de jury, pour l’intérêt qu’ils ont porté à mon travail.

Ces travaux de thèse se sont déroulés au sein du laboratoire LOCIE dont je remercie tout le personnel et particulièrement Benoit STUTZ pour sa volonté à offrir un cadre propice au travail et à l’échange, ce qui a été fort bénéfique dans une période de grands changements. Je remercie également le CEA de l’INES pour m’avoir permis d’occuper leurs locaux et l’ENSTIB pour m’avoir fourni un nouveau cadre de travail. Mais tous ces bureaux auraient été bien différents sans les personnes qui les occupent, merci donc à toutes ces personnes et mes différents co-bureaux pour leur bonne humeur.

Je tiens également à remercier quelques personnes qui hors du cadre du travail m’ont suivi et aidé de différentes façon. La Choubi Team pour cette belle expérience qu’a été la colocation et toutes ces pauses café. Jeanne pour son soutien et son aide précieuse durant la rédaction. Mes amis de Marseille et particulièrement Fabien et Marion pour leur capacité à traverser la France ne serait-ce que pour une demi-journée. Et bien évidemment ma famille qui même à distance reste proche de moi et me soutient.

TABLE DES MATIERES

Remerciement	i
Table des matières	iii
Nomenclature	vii
Introduction	xi
Chapitre I : La réhabilitation et ses problématiques.....	1
1 Contexte énergétique du bâtiment.....	1
1.1 L'évolution réglementaire de la construction	1
1.2 Une meilleure gestion du parc pour réduire les consommations.....	3
1.3 Les niveaux d'exigence actuels pour la réhabilitation.....	5
2 Objectif de la réhabilitation – Une démarche de performance	8
2.1 Etudier les bâtiments pour fixer des objectifs variables	8
2.2 La recherche de performance par la simulation thermique dynamique	9
2.3 Aspect multicritère des exigences du bâtiment d'habitation	12
3 Problématique de l'optimisation : automatique vs transparence.....	14
3.1 Comment trouver la meilleure solution ?	15
3.2 Limitation – La solution est-elle l'optimisation ?	18
3.3 Vers plus de transparence.....	19
4 Indicateurs de performance	21
4.1 Méthodologie de la création des indicateurs	21
4.2 Les principaux points d'évaluation.....	22
5 Un besoin de méthode innovante pour la recherche de solution en réhabilitation	34

Chapitre II : Outils pour l'analyse du comportement des modèles thermiques	37
1 L'analyse de sensibilité pour le bâtiment	37
1.1 Les objectifs de l'analyse de sensibilité	37
1.2 Choix du type de méthode d'analyse de sensibilité pour l'étude des bâtiments	38
2 Théorie des techniques d'analyse employées	40
2.1 Description de la décomposition de la variance – ANOVA	40
2.2 Evaluation de la variance – Méthode de SOBOL	45
2.3 Techniques fréquentielles – FAST	48
3 Améliorations pour la réduction du coût de calcul – RBD	53
3.1 SOBOL' et améliorations.....	55
3.2 FAST et améliorations	57
3.3 Estimateur de confiance – Bootstrap	62
4 Échantillonnage des entrées	64
4.1 L'importance du choix des évaluations	64
4.2 Définition statistique des entrées.....	64
4.3 Les méthodes d'échantillonnage	66
5 Un besoin d'intégration de l'analyse de sensibilité pour l'étude des modèles thermiques.....	70
Chapitre III : Intégration dans une démarche de compréhension du modèle	73
1 Le programme.....	75
1.1 Plan général du programme.....	75
1.2 Définition des entrées et génération des fichiers de simulation	77
1.3 Réalisation des simulations	82
1.4 Extraction et traitement des sorties	84

2	Exemple d'utilisation : Etude d'éclairage sur un modèle simple de bâtiment.	88
2.1	Définition du cas d'étude	88
2.2	Analyse du comportement	90
2.3	Calcul de l'incertitude	94
2.4	Bilan de l'utilisation du programme	96
3	Evaluation multicritère du comportement d'un bâtiment réel	96
3.1	Définition du cas d'étude	97
3.2	Etude des indices de sensibilité	100
3.3	Evaluation de la solution retenue	111
3.4	Bilan de la méthodologie	113
4	Une première étude encourageante mais des coûts de calcul trop important	114
Chapitre IV : Création d'un méta-modèle pour la recherche de solution		117
1	Objectifs du méta-modèle	117
2	Le chaos polynomial	120
2.1	Principe de la décomposition en chaos polynomial – PCE	120
2.2	Exemple de décomposition simple et lien avec les indices de sensibilité.	123
2.3	Calcul des coefficients des polynômes	125
2.4	Indicateur d'erreur	128
3	Comparaisons avec le modèle réel	129
3.1	Calcul d'erreur et convergence	130
3.2	Comparaisons qualitatives	132
3.3	Indices de sensibilité	134
4	Utilisation du méta-modèle – Interface utilisateur	137
4.1	Présentation de l'interface graphique	137

4.2	Lecture des données	139
5	Le méta-modèle une réponse aux contraintes de l'étude thermique	144
Conclusion	147
Perspectives	151
Bibliographie	153
Résumé	163
Abstract	165

NOMENCLATURE

1 Abréviations

Notation	Signification
BBC	Bâtiments Basse Consommation
CPU	Central Processing Unit
FAST	Fourier Amplitude Sensitivity Test
HPE	Haute Performance Energétique
LHS	Latin Hypercube Sampling
NaN	Not a Number
PE	Peu Emissif
RBD	Random Balance Design
STD	Simulation Thermique Dynamique

2 Lettres majuscules (sans unité ou dépendante du contexte)

Notation	Signification
A, B	Matrices d'échantillonnage composées de différentes évaluations de \mathbf{x}
$\mathbf{A}^{(n)}$	n-ème ligne de la matrice d'échantillonnage
$\mathbf{A}(\sigma_i)$	Matrice d'échantillonnage permutée
\mathbf{B}_i	Matrice d'échantillonnage modifiée à la i-ème colonne
C_m^i	Coefficient binomial
$D_{\mathbf{x}}$	Domaine de variation de \mathbf{x}
N	Nombre de simulations
M	Facteur d'interférences
\mathcal{M}	Modèle de simulation thermique dynamique
\mathcal{M}_{PC}	Modèle obtenu par chaos polynomial
S_i	Indice de Sobol du premier ordre
S_{ij}	Indice de Sobol par interaction
ST_i	Indice de sensibilité total
V	Variance totale

Notation	Signification
V_i	Variance dues uniquement à la i-ème variable
V_{ij}	Variance d'interaction entre la i-ème et la j-ème variable
\mathbf{Y}	Vecteur composé de différentes évaluations du modèle

3 Opérateurs et fonctions

Notation	Signification
$E(\cdot)$	Espérance mathématique
$G_i(\cdot)$	Fonction de répartition
p_x	Fonction de densité de probabilité de \mathbf{x}
$V(\cdot)$	Variance
$Y \mathbf{x}$	Y conditionné à une valeur \mathbf{x}
$\hat{\cdot}$	Estimateur
$\bar{\cdot}$	Moyenne

4 Lettres minuscules (sans unité ou dépendante du contexte)

Notation	Signification
a	Constante à définir
a_i, b_i, c_i	Coefficients de Fourier : réel, imaginaire, complexe
f_0	Moyenne du modèle
f_i	Effet seul de la i-ème variable
f_{ij}	Effet d'interaction entre la i-ème et la j-ème variable
k	Nombre de sorties étudiées
m	Nombre de variables d'entrée
\mathbf{x}	Vecteur composé des entrées du modèle
x_i	i-ème entrée du modèle
y_i	i-ème sortie du modèle

5 Lettres grecques (sans unité ou dépendante du contexte sauf si spécifiée)

Notation	Signification
δ	Symbole de Kronecker
ε	Erreur
μ	Moyenne
σ	Ecart type
σ_i	Permutation circulaire
$\phi_i^{(j)}(x_i)$	i-ème polynômes de la base correspondante à la j-ème variable
$\psi_i(\mathbf{x})$	i-ème vecteur de la base de polynômes multi-variés
ω	Pulsation [s ⁻¹]

6 Indices

Notation	Signification
i	Relatif à la i-ème variable
$-i$	Relatif à toutes les variables sauf la i-ème variable

Chapitre d'équation 1 Section 1

INTRODUCTION

La réduction des consommations énergétiques est un défi majeur de notre temps. La communauté scientifique du bâtiment porte pour cela une constante attention au développement des simulations numériques. Les codes de simulation permettent de détailler de nombreux phénomènes relatifs aux équipements et à l'environnement interne du bâtiment tout en donnant des résultats cohérents. Le regroupement de ces outils autour de plateformes de simulation appréhende le système dans sa globalité et facilite la conception de bâtiments à haute efficacité. Le secteur du bâtiment utilise alors ces outils pour répondre à des objectifs de performance de plus en plus ambitieux. Cette performance est le fruit d'une conception et une mise en œuvre soignée et permet au final une réduction des consommations.

Le secteur de la réhabilitation est particulièrement sensible à la problématique de la réduction des consommations énergétiques. Pour ce type de bâtiment, les simulations thermiques permettent d'une part d'identifier les voies d'amélioration et d'autre part de justifier l'opération pour la recherche de financement.

Dans le milieu de la recherche, cette démarche de réhabilitation passe actuellement par des techniques d'optimisation afin d'aider le concepteur dans la recherche et le choix de solutions. Mais ces techniques sont peu transparentes pour l'utilisateur, ne peuvent être employées que dans les cas où il est possible de retranscrire l'objectif en termes de performance de façon précise et ne sont de ce fait que très rarement appliquées.

D'autre part, les simulations thermiques du bâtiment sont déterministes dans le sens où une valeur unique de performance est estimée dans le résultat de simulation. Les valeurs d'entrée du modèle sont fixées et supposées représentatives de la réalité. Or chaque entrée est entachée d'une incertitude et la variabilité des paramètres est l'une des raisons majeures de la difficulté à prévoir le comportement des bâtiments. Elle peut être intrinsèque au matériau avec une variabilité spatiale et temporelle, mais peut également être relative à la mise en œuvre ou à une dégradation locale. Cette incertitude est alors particulièrement présente en réhabilitation et difficile à évaluer.

Cependant les exigences sur la réhabilitation seront certainement renforcées dans les années à venir, il est alors nécessaire de réaliser des études de qualité, pour être proche de la réalité, et d'estimer leur fiabilité, pour limiter ou au moins évaluer

l'incertitude associée au résultat. Le but premier de ce travail de thèse est la mise en place d'outils afin de mieux appréhender la simulation thermique dynamique et de proposer des solutions de réhabilitation pour pouvoir répondre aux problématiques environnementales actuelles tout en prenant en compte les incertitudes.

De nos jours, ces outils de simulation ont acquis une certaine maturité et les stations de travail disposent de puissances de calcul importantes. Le domaine de l'analyse de sensibilité apporte des solutions innovantes pour étudier le modèle thermique, l'analyser et déterminer les objectifs que pourra atteindre le bâtiment vis-à-vis de son état initial. Elles permettront d'appréhender les simulations de façon différente et de se servir de la variation des entrées sur de grandes plages afin de rechercher les valeurs idéales menant à l'optimisation des capacités du bâtiment existant. Ces techniques aideront également à évaluer les causes probables d'un écart entre simulation et réalité en appliquant aux entrées des variations relatives aux incertitudes réelles.

Ce travail propose dans un premier temps, un outil complet pour réaliser automatiquement diverses analyses sur les relations entre les entrées et les sorties du modèle. Les résultats amènent des informations importantes sur le comportement du bâtiment lors de la modification de sa composition à travers la variation des entrées considérées. Les différentes solutions de réhabilitation sont envisagées dans une méthodologie basée sur l'étude de ces relations.

Dans un second temps, cette méthodologie est appliquée dans une interface graphique. Celle-ci utilise un méta-modèle issu d'une décomposition en polynômes faite à partir des simulations obtenues avec le modèle thermique dynamique initial. Ce nouveau modèle permet de réaliser des évaluations pour toutes les configurations envisageables sans surcoût de calcul et avec une précision équivalente. Son intégration dans une interface réactive produit un outil qui permet de sonder le comportement des sorties et offre à l'utilisateur une vision claire des principales relations présentes dans le modèle. Une évaluation instantanée de l'incertitude est possible en tout point de l'espace de variation des entrées et permet ainsi de vérifier la robustesse des choix réalisés.

Les travaux présentés s'articulent de la manière suivante.

Le Chapitre I commence par la description du contexte de la réhabilitation pour cibler les objectifs qui vont être définis dans les prochaines années et ainsi les raisons qui poussent à réaliser des études énergétiques détaillées des bâtiments. Une seconde partie traite de la recherche de performance qui passe alors par la modélisation puis

l'évaluation de plusieurs paramètres de sortie. Ensuite, les possibilités d'arbitrage de ces différentes sorties sont présentées. Celles-ci sont délicates à exprimer et peuvent difficilement prendre une forme générique. Les méthodes d'optimisation sont donc exclues et un autre moyen d'évaluer le bâtiment basé sur l'analyse de sensibilité est mis en place. Enfin une dernière section détaille la méthodologie d'évaluation de la performance et les sorties d'évaluation considérées.

Le Chapitre II détaille les méthodes d'analyse de sensibilité qui seront mises en place. Il est tout d'abord question de voir comment celles-ci permettront l'amélioration des performances et de déterminer le type de méthodes qui seront adaptées à la simulation thermique dynamique. Dans un second temps, la famille de méthodes basée sur l'analyse de la variance est choisie et le principe de deux techniques est détaillé. Dans une troisième partie, les améliorations récentes dont elles disposent seront présentées afin de réduire les coûts de calcul par rapport aux informations demandées. Une évaluation de l'erreur est alors proposée afin de quantifier la qualité des résultats vis-à-vis du nombre réduit de simulations. Enfin des méthodes d'échantillonnages sont décrites pour assurer une bonne utilisation de ces techniques et pour réaliser une analyse d'incertitude.

Le Chapitre III porte sur l'outil développé pour appliquer rapidement les méthodes précédemment décrites. Tout d'abord, le principe général du programme est présenté et les principales étapes, de la définition des fichiers jusqu'à la mise en forme des sorties, sont détaillées. Celui-ci est dans un premier temps appliqué à un modèle simple, mais disposant de modifications complexes. Dans un second temps, une configuration réelle est évaluée et les résultats d'analyse de sensibilité de l'étude sont analysés pour définir la solution optimale. Les informations fournies par le programme donnent alors assez d'informations à l'utilisateur pour réaliser un choix logique de valeurs pour la réhabilitation.

Le Chapitre IV permet d'accélérer cette recherche. Définir et valider le choix de réhabilitation peut demander de nombreuses études. L'utilisation d'un méta-modèle permet d'optimiser les simulations obtenues avec le précédent programme, pour analyser son comportement *a posteriori*. La théorie relative à la construction d'un méta-modèle par une décomposition en chaos polynomial est alors détaillée. Ce nouveau modèle est ensuite comparé avec le modèle initial pour valider les estimations faites. Celui-ci est alors intégré dans une interface qui permet à l'utilisateur de visualiser les résultats d'analyse et le comportement du modèle à la fois pour la recherche de solution et pour la validation quant aux incertitudes.

Chapitre I : LA REHABILITATION ET SES PROBLEMATIQUES

1 Contexte énergétique du bâtiment

1.1 L'évolution réglementaire de la construction

Dans sa définition actuelle, un logement est une unité d'habitation construite pour assurer d'un point de vue architectural plusieurs fonctions. Il doit offrir une protection mécanique et thermique tout en ayant une géométrie et une esthétique qui lui permet de s'intégrer à son environnement. Il permet, de plus, d'assurer à ses occupants des conditions de vie, de sécurité et de salubrité tout en limitant les consommations d'énergie (Larceneux, 2010).

Au milieu du siècle dernier, cette définition était différente, les habitations étaient confrontées à des problématiques de sécurité et de salubrité. De gros travaux ont été réalisés pour assurer ces fonctions et de nos jours les problématiques structurelles et sanitaires sont des prérequis. Cela a permis de redéfinir la notion de confort telle qu'on la connaît actuellement.

L'aspect énergétique dans la définition des bâtiments est arrivé avec les premiers chocs pétroliers après 1973. L'augmentation permanente du prix des énergies a fait apparaître la problématique de la dépendance énergétique et les réglementations thermiques ont alors vu le jour afin de limiter les consommations des bâtiments français. Depuis maintenant dix ans ces réglementations se sont enchaînées de façon assez régulière –1974, 1982, 1988, 2000, 2005, 2012– afin d'augmenter les performances thermiques des bâtiments neufs. Depuis 1988, elles sont accompagnées de labels énergétiques mis en place par le gouvernement ou par des groupements de professionnels. Il y a eu une généralisation de ces labels qui ont un effet moteur en définissant des objectifs plus contraignants que ceux de la réglementation en vigueur. Ils ont montré qu'il était possible d'aller au-delà de la réglementation, ont assuré les transitions entre celles-ci et ont permis de valoriser les démarches environnementales

et de recherche de performance. Durant cette évolution est alors apparue une prise de conscience de l'épuisement des ressources et une incitation à la sobriété et à l'utilisation d'énergies renouvelables.

Cette démarche de transition énergétique s'intègre à tous les niveaux de la société : mondial avec *les accords de Kyoto*, Européen avec la mise en place de directives et Français avec les *grenelles environnement*. Il y a un consensus sur la réalité d'un changement climatique et il a été montré avec un niveau de confiance à 95% que nos émissions de gaz à effet de serre (GES) depuis le milieu du 20^e siècle en sont la cause principale (IPCC, 2013). La lutte contre le changement climatique est une priorité majeure qui impose de réduire les émissions de gaz à effet de serre et de mettre en place des démarches pour les années à venir (CCNUCC, 1994). Les mesures nationales imposent ou à défaut incitent à une meilleure efficacité énergétique, à la baisse de la consommation et à la réduction des émissions des grands secteurs de consommation que sont les bâtiments, les transports et les industries. Le secteur du bâtiment est placé au premier rang (Cheng et Pouffary, 2008), il représente près de 40% des dépenses en énergie primaire en Europe, mais a la possibilité de mettre en place des projets à fort impact environnemental avec un investissement relativement faible par rapport aux autres grands secteurs consommateurs. Il dispose notamment d'un large panel de technologies déjà commercialisées pouvant réduire les besoins énergétiques. Le coût de mise en place de ces technologies est alors faible par rapport aux coûts que pourrait engendrer la mise au point et la commercialisation d'une nouvelle technologie ayant le même impact dans le secteur automobile.

Les exigences actuelles proviennent de l'accélération des réformes incitées notamment par le premier Grenelle (2009) qui est lui-même dans la continuité des engagements européens. Parmi les objectifs, on observe notamment la réduction de la consommation des logements neufs sous le seuil des 50 kWh/m².an avant 2012 ce qui a été appliqué avec la RT 2012. Avec le temps, la politique de limitation des consommations a évolué vers une politique de réduction énergétique.

Cet effort va se poursuivre, car de nouveaux objectifs sur le long terme ont été définis avec les projets de loi sur la transition énergétique. Des exigences vont être fixées en matière de réduction des émissions de gaz à effet de serre et d'amélioration de l'efficacité énergétique avec « *une réduction de 50 % de la consommation d'énergie à horizon 2050, de réduction de 30 % de la consommation de combustible fossile à l'horizon 2030 et de diversification du mix électrique avec le développement des énergies renouvelables et la réduction de la part du nucléaire à 50 % à l'horizon 2025* » (Ministère Français de l'Ecologie, 2013). Le parlement européen a également publié un ensemble de

directives à l'attention des pays membres. Dans ces textes, les différentes nations doivent veiller à ce que la consommation d'énergie des nouveaux bâtiments soit quasi nulle dès 2019 pour les bâtiments occupés et possédés par les autorités publiques et dès 2021 pour l'ensemble des bâtiments (Directive 2010/31/UE, 2010). Les textes stipulent également l'obligation pour les États membres d'élaborer des politiques et de définir des objectifs afin de stimuler la transformation de bâtiments rénovés en bâtiments dont la consommation d'énergie est quasi nulle.

C'est sur ce dernier point que nous portons notre attention. Transformer le parc de logements en bâtiments dont la consommation d'énergie est quasi nulle est un objectif très ambitieux qui nécessite la mise en place de méthodologies. Nous verrons dans les parties suivantes, les moyens actuellement mis en place, les outils disponibles et comment faire pour atteindre de tels objectifs.

Toutes les évolutions réglementaires précédentes et les projets d'évolutions futures montrent que le Bâtiment Basse Consommation doit être le nouveau standard, car ce type de bâtiments permet de limiter efficacement les déperditions. Dans ce type de bâtiment, le principal poste de déperdition qui est la ventilation dispose déjà de systèmes de récupération de chaleur et il sera très compliqué de réduire encore ces pertes. De nombreuses applications montrent que des bâtiments zéro énergie ou à énergie positive peuvent dès à présent être construits avec une mise en œuvre soignée et la mise en place de systèmes performants et générateurs d'énergie.

L'évolution de la construction se tourne alors vers une vision plus globale des consommations avec la prise en compte de tous les postes de dépenses et une intégration des émissions de GES. Les pronostics pour les prochains objectifs réglementaires semblent également s'orienter sur l'évaluation des consommations réelles ce qui demande l'instrumentation des logements lors des premières années. On passerait alors d'un objectif défini en matière de moyens à un objectif de résultats. Ce dernier point semble être dans la continuité des mesures d'étanchéité imposées dans la RT2012 et ouvre la voie vers la garantie de performance que nous aborderons plus tard.

1.2 Une meilleure gestion du parc pour réduire les consommations

Les engagements définis précédemment portent sur le long terme et des efforts soutenus sont à réaliser afin d'atteindre ces objectifs. Cependant, les évolutions réglementaires faites jusqu'à présent ne permettront pas de réduire les dépenses énergétiques ni les émissions de GES. Elles assurent cependant une stabilisation des dépenses, définissent un niveau de performance à atteindre et permettent de

modifier l'état d'esprit du secteur, c'est un prérequis pour réduire la facture énergétique.

Le gouvernement français met à disposition les chiffres du logement dont les données sont actualisées chaque troisième trimestre avec les bilans de l'année précédente. Pour 2009, le parc immobilier français comporte 33,1 millions de logements et a subi une hausse annuelle constante de 1,2% sur les 20 dernières années. Le nombre de permis de construire autorisés pour cette année 2009 s'élève à 348 000 ce qui correspond à la croissance annuelle. On peut ainsi conclure que très peu de logements sont détruits. Le renouvellement des logements, qui est généralement estimé à 1%, ne correspond en fait qu'à l'augmentation du parc. Il est alors difficile de parler de renouvellement quand les logements les plus consommateurs ne sont pas remplacés par des logements neufs.

La Figure I-1 montre la répartition des logements par année de construction entre 2001 et 2011 et on observe l'évolution presque inexistante du parc de logement construit avant 1998 quel que soit la période de construction. La seule évolution des réglementations sur le neuf n'est donc pas suffisante, car elle ne concerne qu'une faible part des bâtiments. Il est nécessaire de se tourner vers le secteur de la réhabilitation pour espérer réduire les consommations –malgré l'augmentation du nombre de logements et le développement des cités pavillonnaires.

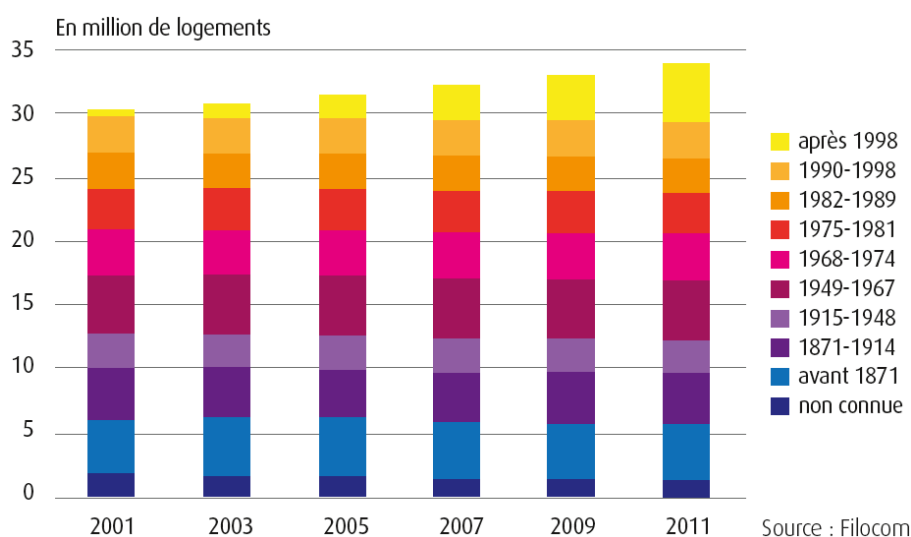


Figure I-1 : Répartition des logements par année de construction

Le graphique précédent montre également que 60% des bâtiments ont été conçus à une période où aucune réglementation thermique n'était présente et de nombreux autres bâtiments ont été construits avec des critères de performances bien loin des standards actuels. « En 2008, la France compte plus de 32 millions de bâtiments dont plus de la moitié ont été construits avant 1975. La consommation moyenne annuelle d'un bâtiment

est d'environ 240 kWh d'énergie primaire par m² chauffé, alors que les enjeux énergétiques et climatiques exigent de réduire cette consommation à une centaine de kWh/m².an. Les potentiels d'économies d'énergie sur le parc bâti sont donc colossaux et exigent une action forte et prioritaire qui doit s'inscrire dans une perspective de long terme » (ADEME).

Ce secteur de la réhabilitation commence à évoluer et des initiatives apparaissent, cependant le parc résidentiel français subit une mutation faite au cas par cas et peu contrôlée. Mis à part les 49 000 demandes de travaux de réhabilitation lourde en 2009, qui ne représentent par ailleurs que 0,1% du parc, toutes les réhabilitations légères issues du secteur de l'entretien-amélioration ne sont pas quantifiables, car elles n'exigent pas de permis de construire. On sait cependant que le chiffre d'affaires des entreprises du bâtiment se répartit de façon plus ou moins égale entre le neuf et la rénovation. Il y a donc de nombreuses opérations réalisées dans les bâtiments existants, mais sans réelle possibilité de référencement. Dans le cas où tous les éléments sont révisés, maintenus et remplacés, cela est fait par des entreprises différentes et sans approche globale. Ces opérations sont étudiées pour assurer une rentabilité à plus court terme possible et n'entre pas dans une réelle démarche de performance vis-à-vis du bâtiment.

Le gouvernement fait de la transition énergétique une de ses priorités. Nous avons vu dans la partie précédente que ce secteur sera amené à se développer dans les années à venir. Le respect des engagements de réduction de GES prévoit ainsi la réhabilitation de 800 000 logements sociaux, dont la consommation annuelle d'énergie est supérieure à 230 kWh/m².an d'énergie primaire, pour réduire leur consommation à 100 kWh/m².an. Mais cet objectif, certes d'envergure, n'est qu'une étape dans la modification du parc. En effet, la quantité de logements qui compose le parc et son évolution induit un important nombre de logements anciens. On estime que 80% du parc de logement européen de 2030 est déjà construit (EC Adhoc Industrial Advisory Group, 2013). Les initiatives sur la mise en œuvre de politiques et de démarches globales de réhabilitations qui seront faites dans les années à venir conditionneront donc de façon significative notre capacité à atteindre les futurs objectifs de consommation. Et d'ici 2030, il est certain que la problématique énergétique sera toujours plus présente.

1.3 Les niveaux d'exigence actuels pour la réhabilitation

Il est donc clair qu'une démarche de réhabilitation qui vise l'amélioration du bâtiment tel que l'Etat le définit doit nécessairement viser des objectifs de haute

performance à long terme. Il est alors nécessaire de différencier dès à présent réhabilitation et rénovation.

La rénovation correspond à une remise à neuf des éléments. Avec l'évolution des technologies et des techniques, il est naturel et parfois plus économique d'employer des éléments couramment utilisés dans le neuf. L'opération a alors de grandes chances de produire une amélioration des performances initiales. Cette démarche est réglementairement imposée depuis 2008 via la RT2005Ex élément par élément. Celle-ci impose, pour toute opération sur un bâti existant, des caractéristiques minimales à respecter sur les éléments visés.

La réhabilitation pourrait être définie par le second volet de cette même réglementation dite RT2005Ex globale. Celle-ci s'applique aux opérations de rénovation lourdes dont :

- la Surface Hors Œuvre Nette (SHON) est supérieure à 1000m²,
- la date d'achèvement du bâtiment est postérieure au 1er janvier 1948
- le coût de la partie thermique des travaux de rénovation étalés sur deux années est supérieur à :
 - 322 €HT/m² pour les logements
 - 275 €HT/m² pour locaux non résidentiels.

Elle impose une étude thermique du bâtiment pour trouver les meilleures voies d'amélioration, une analyse des solutions d'approvisionnement en énergie et fixe également une limite de consommation des postes chauffage, refroidissement et ECS de 80 à 165 kWh/m².an en fonction des cas. Cette partie de la réglementation est importante et est associée à une réelle réflexion sur le bâtiment. Cependant elle ne correspond qu'à une petite partie des demandes de travaux et comme nous l'avons vu précédemment le nombre d'opérations de ce type est faible –pour rappel le nombre total de demandes de travaux de réhabilitation par an est de l'ordre de 0,1% du parc. Cela peut s'expliquer notamment par la nécessité d'intervenir en site inoccupé, mais également par la difficulté d'effectuer une étude thermique à partir d'un bâtiment existant et rechercher des améliorations.

Actuellement, le cadre réglementaire sur le domaine du logement existant est limité. Mais, il est certain que ces réglementations évolueront dans les années à venir. Lors du *Grenelle 2*, certains textes originaux, qui n'ont pas été validés, précisaient la nécessité de se mettre en accord avec la réglementation thermique et acoustique pour toute opération nécessitant la demande d'un permis de construire à partir de 2012. Ce point de réglementation est de nouveau étudié dans le projet de loi sur la

transition énergétique : « Concernant la rénovation thermique des logements une obligation pourrait notamment être introduite au travers d'une modification de la réglementation thermique existante, avec un dispositif de prise en compte de l'efficacité énergétique lors de moments importants de la vie du bâtiment, réfection de toiture et ravalement en particulier. Concernant les bâtiments tertiaires, suite à l'obligation de rénovation à 2020 figurant dans la loi Grenelle et dont le décret d'application paraîtra en 2014, une visibilité à plus long terme sera définie ». (Ministère Français de l'Ecologie, 2013)

Les labels mis en place pour le parc immobilier existant peuvent nous donner une idée de ce que pourrait être la réglementation dans les années à venir. On remarque ainsi qu'il est possible d'atteindre des performances niveau BBC en réhabilitation. Le tableau suivant présente les exigences sur les consommations définies en énergie primaire et comprenant chauffage, ECS, refroidissement, éclairage, auxiliaires de chauffage et de ventilation – ainsi que les équipements pour le Passiv Haus.

Tableau I-1 : Récapitulatif des exigences de quelques labels en rénovation
(Conseil d'Architecture d'Urbanisme et de l'Environnement)

	Consommation maximum (kWh/m²/an)	Autres exigences
HPE rénovation 2009	~150 selon climat	Confort d'été
BBC rénovation 2009	~80 selon climat	Confort d'été
Effinergie rénovation	~80 selon climat	Perméabilité à l'air
Minergie	60	Récupération de chaleur
Passiv Haus	120 (42 en énergie finale dont 15 pour le chauffage)	Perméabilité à l'air Suppression des ponts thermiques Triple vitrages

Il faut cependant noter que demander une mise en conformité complète est une demande administrative lourde pour toute opération et devient un frein pour les propriétaires. La définition d'une nouvelle réglementation est très délicate et nécessite des règles applicables sur un ensemble très disparate d'opérations. Le gouvernement incite alors principalement à la réhabilitation par des avantages par des crédits d'impôt ou une réduction de la TVA. Mais comme vu précédemment, ce type d'opération se limite souvent à la rénovation et donc au remplacement d'un élément au lieu d'une étude globale comme l'exigerait une réhabilitation de l'existant.

Il est donc nécessaire de réaliser une étude du bâtiment afin de définir un niveau de performance qui soit en accord avec les capacités thermiques de celui-ci et également trouver les solutions technologiques et économiques qui permettent d'atteindre ce niveau de performance.

2 Objectif de la réhabilitation – Une démarche de performance

2.1 Etudier les bâtiments pour fixer des objectifs variables

L'évolution des réglementations sur le neuf a mis de nombreuses années à atteindre les performances actuelles. Il a été nécessaire de modifier notre façon de penser et de concevoir le bâtiment. Les objectifs se tournent maintenant vers l'existant et il est question de transférer ce savoir-faire à la réhabilitation. On voit dans certaines opérations et vis-à-vis des labels qu'il est possible d'intégrer dès à présent des objectifs ambitieux. Cela demande cependant de travailler sur la définition de ces objectifs et d'envisager une méthodologie à mettre en place pour les atteindre.

Les objectifs que doit atteindre le bâtiment réhabilité nécessitent d'être adaptés au cas étudié, car il est peu probable de toujours atteindre des performances égales aux contraintes réglementaires actuelles dans le neuf et les objectifs vont dépendre du bâtiment existant considéré. La performance des bâtiments neufs est permise par une intégration dès la phase de conception des problématiques thermiques afin notamment d'appliquer des principes bioclimatiques et optimiser les systèmes énergétiques. C'est cette démarche commune qui permet de contrôler et de minimiser les déperditions pour arriver à des performances élevées.

Dans le cas d'un bâtiment à réhabiliter la structure est existante, elle n'intègre pas forcément les problématiques thermiques notamment l'aspect bioclimatique conditionné entre autres par l'orientation du bâtiment et, le positionnement et le dimensionnement des fenêtres. De plus, certaines opérations, comme l'isolation du plancher, ne sont pas toujours réalisables. Un objectif de performance arbitraire ne peut donc être défini à moins de faire abstraction de l'état initial. Pour tirer le meilleur parti de chaque bâtiment, il est nécessaire de réaliser une étude détaillée afin de définir leur potentiel d'amélioration. Cette démarche permet également d'évaluer le potentiel de gain énergétique sur un bâtiment existant et amène à une prise de décision sur une démolition ou non du bâti en fonction des possibilités ou non d'amélioration énergétique.

Cependant, il est difficile d'étudier un bâtiment pour trouver des solutions quand l'investissement est limité. Un projet est généralement impulsé par l'apparition d'un

désordre ou un état général qui pousse à envisager une remise en état. Cette opération est alors une opportunité pour modifier les caractéristiques du bâtiment afin qu'il réponde mieux aux attentes des occupants et aux contraintes environnementales. Une étude est alors réalisée spécialement pour cette opération, elle se compose d'une étude thermique, d'une étude de faisabilité et également d'une étude économique. Rechercher une solution performante et adaptée est alors complexe quand l'opération a un objectif précis et un budget limité. Et on produit donc une solution à court terme.

Pour atteindre efficacement des performances élevées, il est nécessaire de favoriser les études en amont de tout projet. Généralement, tous les bâtiments subissent au cours de leur fonctionnement plusieurs opérations de gros entretiens qui peuvent facilement être anticipées. Pour que ces opérations puissent améliorer progressivement les performances du bâtiment, il est nécessaire de les prévoir afin de rechercher la meilleure façon d'intégrer l'objectif énergétique qui était généralement absent lors de la construction.

Le Diagnostic de Performance Energétique est un début de réponse, c'est une évaluation grossière des capacités du bâtiment vis-à-vis des standards actuels afin de donner une information sur l'état énergétique du bien. Des cartographies thermiques sont également réalisées afin d'évaluer l'état des toitures ou des façades. Toujours dans cette idée, mais outre-Atlantique, l'état de Californie impose via les textes de loi AB758 et AB1103, une modélisation thermique pour tous les bâtiments commerciaux lors des changements de propriétaire (Bhandari et al., 2012).

Réaliser une étude en amont permet de prévoir des investissements et de détecter les problèmes avant qu'ils n'apparaissent. Le bâtiment entre donc naturellement dans une démarche de suivi. La première étape est donc l'étude du bâtiment pour déterminer quel peut être le niveau de performance atteignable.

2.2 La recherche de performance par la simulation thermique dynamique

Pour analyser le bâtiment, il est nécessaire de se doter de méthodes d'évaluation. C'est-à-dire un outil qui intègre l'ensemble des problématiques thermiques et qui permet la recherche de solutions. Cela nécessite de définir les limites de l'étude, connaître les besoins du calcul et des informations disponibles sur l'objet étudié.

L'objectif est d'identifier les faiblesses du bâtiment ainsi que des pistes pour avoir une réduction ou une meilleure gestion des échanges énergétiques, ceux-ci jouant à

la fois sur les conditions intérieures et les consommations. Il est donc nécessaire d'avoir accès au comportement détaillé du bâtiment. Les conditions intérieures sont importantes pour plusieurs raisons, tout d'abord la réduction des consommations ne peut se faire au détriment du confort d'autant plus que lors de la réhabilitation l'amélioration du confort d'hiver peut être un élément important. Et d'autre part, sachant que l'état final peut être un niveau BBC, des problématiques de surchauffe estivale peuvent apparaître. Il est ainsi important d'évaluer les températures dans un fonctionnement en évolution libre.

L'outil doit s'adapter pour évaluer tous type de bâtiments, neufs ou existants et doit intégrer aussi bien l'étude de l'état initial que celle de l'état final du bâtiment étudié. Il doit pouvoir réaliser des simulations fines sur des périodes particulières et modulables afin de réduire les coûts de calcul et de permettre une phase de dimensionnement plutôt que des vérifications réglementaires. D'autre part, les calculs de consommation tendent à intégrer toutes les dépenses énergétiques du bâtiment notamment le fonctionnement des systèmes ou des éclairages. Ces systèmes disposent de plus en plus de mécanismes de régulation complexes. Permettre leur prise en compte est alors un avantage non négligeable d'autant plus que la validation réglementaire peut demander ce type d'étude.

C'est pourquoi il a été choisi de réaliser des Simulations Thermiques Dynamiques (STD) qui permettent l'intégration au pas de temps horaire ou inférieur d'une multitude de systèmes et de conditions limites en interaction avec le bâtiment. De plus, les outils de STD se généralisent dans les bureaux d'étude. Ils servent pour la conception, la validation des projets et les outils réglementaires tendent à s'en inspirer.

En contrepartie, ces outils demandent un plus grand investissement de la part du modélisateur. La prise en main est parfois laborieuse et de nombreuses données d'entrée sont nécessaires pour profiter pleinement des capacités. Une représentation plus détaillée du bâtiment permet un meilleur contrôle de la physique du modèle, car la modélisation physique est plus proche de la réalité. Mais ce n'est pas parce qu'un résultat est plus détaillé et donc de meilleure qualité qu'il sera plus précis. Pour évaluer la précision, il est nécessaire de retranscrire correctement le bâtiment dans le modèle et intégrer l'aspect incertain des entrées qui peut être pour certaines significatif. On peut notamment penser à l'aérodynamique ou à l'hygrométrie du bâtiment.

Cette problématique ne dépend donc pas complètement des outils. La première étape d'une étude est la collecte d'information, car il est courant que les documents du projet initial ne soient plus disponibles ou dans le cas contraire, que les caractéristiques des matériaux de construction aient changé. Il est ainsi nécessaire de mettre en place des moyens et méthodes de caractérisation des éléments in situ –ou lors de la mise en œuvre pour évaluer une opération. Les questions à se poser sont méthodologiques et techniques. Il est nécessaire de dresser une liste des points à inspecter pour avoir une vue complète du logement et ne pas oublier d'éléments. Il faut déterminer les variables nécessaires à l'étude afin d'éviter de surcharger la collecte d'information par des éléments trop incertains, inutilisables, incomplets ou simplement inutiles.

C'est dans cette optique que le projet JOULE a mis au point la méthode EPIQR (Genre et al., 2000). Elle permet une première évaluation des bâtiments aussi bien sur le point de vue thermique que structural. A partir d'un scénario de visite qui recense les points importants du bâtiment et de questionnaires à l'attention du propriétaire et des occupants (Bluyssen, 2000), cette évaluation permet de dresser un bilan du bâtiment afin de mettre en avant les points faibles. Cette méthode est pratique et fonctionnelle. Elle est généralement utilisée avec des méthodes d'analyse de risque, car les informations recueillies sont le plus souvent de nature qualitative, mais pourrait servir de base pour une méthodologie destinée à la simulation thermique dynamique.

D'autre part, le résultat dépend également de l'utilisateur et de sa connaissance de l'outil. Le passage à une représentation énergétique impose des hypothèses et des choix de modèles qui en fonction des outils limitent les possibilités ou sont source d'erreurs. Ce n'est enfin que lors de l'étape de calibration du modèle qu'il est possible de vérifier les erreurs dues aux parties précédentes. C'est alors que l'analyse de sensibilité peut apporter des pistes pertinentes d'amélioration, certaines de ces techniques sont présentées dans le Chapitre II. Elles permettent d'analyser le comportement du modèle numérique et offrent un apport important d'informations pour la calibration ou une recherche de solution, que ce soit en réhabilitation ou sur bâtiment neuf.

Au final, la simulation thermique dynamique permet une étude fine du comportement thermique du bâtiment, mais nécessite une définition précise. Elle est cependant la seule à pouvoir apporter tous les éléments nécessaires pour rechercher les voies d'amélioration et valider des solutions innovantes. Des efforts doivent donc être réalisés afin d'apporter les éléments nécessaires à la bonne définition de ces

simulations. Et à défaut, le suivi des bâtiments dans le temps et dans les performances peut permettre d'augmenter progressivement la connaissance du comportement.

Le temps de calcul est un élément souvent mis en avant dans les simulations thermiques dynamiques. Les durées de simulation restent acceptables et sont constamment réduites par l'évolution des capacités de calcul. D'autre part il est toujours possible de simplifier les modèles pour produire des simulations dont le temps de calcul est alors inférieur à 20 secondes. Un ordinateur personnel actuel avec 4 cœurs peut alors réaliser 1000 simulations en 1h30. Une recherche de solution est alors possible même sans disposer d'une importante puissance de calcul.

2.3 Aspect multicritère des exigences du bâtiment d'habitation

La simulation thermique dynamique est alors employée pour réaliser une recherche de performance. Pour l'instant, la principale performance qui a été définie et qui est au cœur des réglementations est la réduction de la consommation, mais la littérature fait état de différentes recherches de performance en fonction de l'orientation que l'on souhaite donner à l'amélioration du bâtiment. Généralement, des éléments d'évaluation sont choisis en fonction de l'objectif principal qui peut viser les cibles suivantes :

- Investissement (amélioration de l'image, réalisation d'une plus-value)
- Réduction des coûts (maintenance / entretien / fonctionnement)
- Image (économique / écologique / impact environnemental)
- Amélioration de la qualité de vie
- Mise en conformité

Il est nécessaire de s'interroger sur les éléments à considérer pour l'analyse et vérifier si la liste est exhaustive afin d'évaluer correctement la qualité et l'impact de chacune des modifications effectuées. Il est donc important de bien définir ces points, car ils conditionnent la solution à mettre en œuvre pour répondre aux attentes.

Cependant, l'importance attribuée à chacun des points précédents ne sera pas la même en fonction de la personne qui pose le problème et son rôle dans le projet. Il existe plusieurs niveaux de contrainte ou de décision à considérer :

- Politique nationale
- Gestion
- Conception
- Utilisation

De nombreuses personnes gravitent autour de ce système et elles ont des problématiques différentes. Même si la décision finale appartient au gestionnaire, réaliser une opération demande de prendre en considération de nombreux paramètres : Le contexte du bâtiment, les attentes des occupants et de la maîtrise d'ouvrage, les contraintes de site ou technologies et les réglementations. La démarche est ainsi multicritère et la recherche de solution de réhabilitation doit permettre d'estimer le meilleur compromis par rapport aux attentes qui ont été définies, car aucune solution ne pourra améliorer simultanément tous ces points. On peut en conclure qu'une démarche multicritère globale n'est donc pas une optimisation simultanée de tous ces éléments et il n'y a pas d'opération idéale. Une recherche de solution doit être réalisée et les résultats analysés pour trouver le meilleur compromis.

D'autre part, la vision du bâtiment tend à être de plus en plus globale. Ainsi le bâtiment est le fruit de l'interdisciplinarité ce qui implique que les contraintes à prendre en compte lors de la mise en œuvre sont multi-domaines :

- Environnemental
- Financier
- Mécanique
- Sociologique
- Technologique
- Thermique

Tous ces éléments ne sont pas intégrés dans le travail de thèse, mais il a paru intéressant de les mettre en avant pour montrer la complexité de la gestion multicritère du bâti et de sa réhabilitation. C'est ainsi que l'aspect mécanique intervient, car la bonne tenue de la structure est une condition *sine qua non* pour toute démarche de réhabilitation, mais elle ne fait pas partie des variables d'optimisation. On peut également considérer l'aspect sociologique, par exemple, une solution d'économie d'énergie efficace, mais trop complexe dans son utilisation pourra être mise en échec si elle n'est pas intégrée pas les utilisateurs. Il y a également des aspects qui prendront de l'importance avec le temps telle que l'Analyse de Cycle de Vie qui considère notamment l'énergie grise – énergie nécessaire à la fabrication et la mise en œuvre de l'ensemble des matériaux utilisés pour l'opération. C'est un élément important dans la démarche environnementale du bâtiment et dont l'importance augmente avec les performances thermiques et les moyens mis en place.

Pour un bâtiment ancien, l'énergie grise dépensée dans l'ouvrage représente une part relativement faible face aux dépenses énergétiques dues au fonctionnement. Au

contraire, pour un bâtiment plus récent cette part augmente, car les meilleures performances ont généralement demandé une mise en œuvre et des matériaux plus coûteux (Monahan et Powell, 2011). Une opération de réhabilitation va encore augmenter le coût CO2 du bâti et produire une réduction des dépenses associées. L'énergie grise dépensée pour l'opération peut, en fonction des situations, correspondre à de nombreuses années de consommation. Le temps de retour vis-à-vis de l'empreinte carbone peut ainsi être très long et amener à se questionner sur la manière d'intervenir ou même sur l'utilité environnementale de l'opération. Un nouveau compromis peut donc apparaître, ce compromis n'est toutefois pas présenté dans ce travail de thèse.

Au final, de nombreux paramètres peuvent être pris en compte avec les niveaux d'exigences variables selon le projet et dans le temps en fonction de l'évolution des besoins ou des résultats attendus. La recherche d'une solution de réhabilitation est donc une tâche complexe qui nécessite une vision holistique du bâtiment qui dépend également des différents acteurs du bâtiment et de la phase de projet. Il est courant de limiter le nombre de facteurs à cause de la difficulté pour analyser de façon simultanée un ensemble de fonctions d'évaluation antagonistes. La recherche de solution en réhabilitation s'inscrit pleinement dans une problématique multicritère. D'une façon générale plus il y aura de variables à optimiser et donc d'objectifs et plus la recherche de solution sera difficile. Or le but de ces travaux de thèse n'est pas pour l'instant de faire la recherche la plus complexe. Le plus important est semble-t-il de se laisser la possibilité de prendre les facteurs que l'on souhaite en considération en fonction du projet traité.

Quel que soit le nombre de variables d'évaluations choisies, l'amélioration des performances nécessite donc d'analyser le bâtiment et d'arbitrer les différents résultats fournis par ces variables d'évaluations.

3 Problématique de l'optimisation : automatique vs transparence

La simulation thermique dynamique est utilisée pour obtenir une estimation précise du fonctionnement du logement avant travaux. Elle s'impose donc dans la recherche de solution afin de déterminer la solution la plus appropriée pour le cas étudié. Un problème multicritère est considéré et uniquement les variables d'évaluation les plus importantes sont prises en compte. Le processus de décision consiste à rechercher par cet outil le ou les meilleurs résultats dans l'ensemble de solutions envisagées. Ces solutions correspondent aux variables qui seront modifiées au sein de l'outil de simulation. La recherche par tâtonnement devient alors une tâche difficile à cause du

nombre de combinaisons possibles et impose *a priori* l'usage de méthodes d'optimisation. Les parties suivantes porteront donc sur quelques principes de base sur la recherche de solution. Nous verrons cependant que toutes les méthodes nécessitent l'aide du savoir expert, c'est-à-dire l'aide de la connaissance du spécialiste.

3.1 Comment trouver la meilleure solution ?

Les méthodes d'optimisation couvrent un large champ d'application et disposent de nombreuses techniques. L'objectif principal de ces algorithmes est d'identifier le plus rapidement possible une solution idéale sans se soucier du comportement général du modèle. Dans la démarche classique, une variable d'évaluation globale permet de donner une valeur à chaque simulation pour estimer la qualité de la solution. La solution idéale est alors celle qui maximise ou minimise la variable d'évaluation, aussi appelée fonction objective. Pour réaliser cela, un nuage de points est évalué pour permettre à la solution d'évoluer au fil des itérations dans la bonne direction. Un large choix de méthodes est disponible parmi lesquelles on peut citer les Algorithmes Génétiques, Pattern Search Algorithms, Essaims de particules, *etc.*

Les outils d'optimisation ont généralement besoin d'un nombre important d'évaluations et ce nombre augmente avec le nombre de variables d'entrée. L'algorithme doit être assez robuste pour ne pas être affecté par les discontinuités dues au modèle thermique. Les outils de simulation doivent résoudre des équations différentielles complexes et utilisent des solveurs adaptatifs pour obtenir des solutions approximatives. Lors d'une modification des entrées, ces solveurs peuvent modifier leurs itérations ou leurs maillages, ce qui produit des discontinuités dans les sorties (Wetter et Polak, 2005). Un autre problème peut survenir lors de la création de la fonction objective ou variable d'évaluation, celle-ci peut comporter des miniums locaux ou des discontinuités que l'algorithme devra éviter – espace convexe/concave.

La construction de cette fonction objective a évolué au fil des ans. Traditionnellement, les techniques d'analyse monocritère recherchent la minimisation des coûts de fonctionnement et des coûts de conception. Elles ont fourni une aide aux concepteurs afin d'identifier les options les plus avantageuses pour produire à bas prix. Les méthodes ont ensuite suivi une évolution similaire à l'évolution des réglementations qui a été présentée en début de chapitre section 1.1. L'augmentation des besoins de performance des bâtiments et l'arrivée des problématiques environnementales ont orienté l'attention des concepteurs vers les méthodes d'aide à la décision multicritère (Multi-Criteria Decision Making). Dans le

cas d'une recherche multicritère, les deux principales familles de méthodes sont Multi-Attribute Decision Making (MADM) et Multi-Objective Decision Making (MODM) (Pohekar et Ramachandran, 2004)

3.1.1 *Multi-Attribute Decision Making*

Les méthodes MADM permettent de réduire la dimension du problème pour se ramener à une unique sortie d'évaluation ou fonction objective. L'atteinte de l'objectif final est alors déterminée par cette nouvelle sortie. Classiquement, les techniques de ce type agrègent les informations pour former une fonction objective qui permettra de choisir parmi les solutions possibles. Par exemple, le principe de la « méthode des sommes pondérées » est de normaliser les différents éléments puis de leur attribuer un poids afin de former une unique variable à minimiser.

$$g_{sum}(\mathbf{y}) = \sum_{i=1}^k p_i \cdot F_i(\mathbf{y}) \quad (I.1)$$

D'autres méthodes permettent de produire des fonctions plus complexes qu'une simple pondération, mais le principe reste similaire et demande de minimiser un unique résultat. Dans tous les cas, il faut pouvoir produire la fonction objective g_{sum} qui s'adapte parfaitement au problème, car le résultat dépend entièrement de cette fonction. Il y a beaucoup d'outils qui utilisent ces méthodes, mais souvent pour des applications particulièrement bien définies, car la solution est fortement contrainte par la fonction objective et donc des pondérations associées à chaque variable de sortie.

Une extension de ces méthodes consiste en l'ajout d'une variable de limitation en plus de la variable d'évaluation. Par exemple, on recherche la solution qui permet de réduire les consommations tout en écartant les solutions qui ne satisfont pas le critère de confort (Wright et al., 2012). La méthode employée doit être capable d'écarter des solutions sans perturber le fonctionnement de l'algorithme de recherche. Cette méthode est alors généralement associée aux algorithmes génétiques. Cependant, en présence de deux sorties ayant des comportements opposés la solution obtenue est encore contrainte et unique. Elle dépend du poids attribué à chacun des éléments qui composent la fonction objective et/ou dépend de la valeur des critères de limitation choisis. Dans l'exemple précédent, la solution idéale sera à la limite du confort acceptable qui a été défini, et dépend donc fortement de cette valeur (Hamdy et al., 2011).

Garder l'aspect multicritère implique donc de réaliser des compromis. Il peut être réalisé à travers la fonction objective, mais cela demande une part importante de savoir expert, car il est nécessaire de définir ces éléments sans avoir d'information sur la forme des sorties ni sur les performances qu'il est possible d'atteindre. Cela ajoute une contrainte sur la solution qui n'a pas lieu d'être et on réalise ainsi un choix avant même le début de l'analyse en la conditionnant fortement.

Certains travaux proposent ainsi des outils qui évaluent l'impact des pondérations choisies sur l'optimisation cela en utilisant des pondérations aléatoires (Butler et al., 1997). Les poids attribués aux différentes variables de sortie qui servent à la recherche de solution sont évalués afin de déterminer si une pondération différente n'apporte pas une solution globalement meilleure. Cette technique se rapproche ainsi de la seconde famille de méthodes présentée.

3.1.2 Multi-Objective Decision Making

Afin de repousser ce choix, il est possible d'employer des méthodes permettant de faire une recherche MODM. La fonction objective est remplacée par une méthode qui autorise plusieurs solutions optimales. Les solutions sélectionnées sont issues du front de Pareto et sont définies de manière qu'il n'est pas possible d'augmenter une performance sans en dégrader une autre. On obtient une zone de l'espace des sorties qui est optimale.

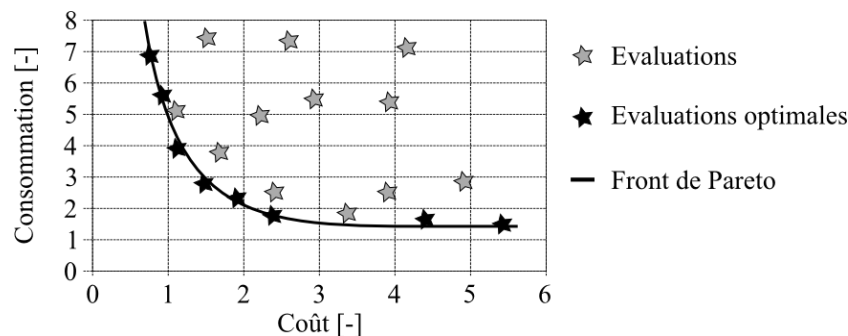


Figure I-2 : illustration du front de Pareto

Dans les méthodes de ce type, on peut citer les méthodes de sur-classement comme ELECTRE qui permettent via des seuils de préférence ou de non-distinction de sélectionner les meilleurs résultats (Ben Mena 2000). Cette technique permet par ces seuils d'intégrer dans la décision qu'une différence de $0,1\text{kWh/m}^2$ sur la consommation n'est pas significative et ne pourra donc pas permettre de faire un choix. Le front de Pareto est élargi et permet de ne pas écarter une solution technique à tort.

La sélection des résultats optimaux est alors non contrainte et la nuance de compromis a été intégrée dans le choix. La solution idéale est accessible, mais une seconde recherche est nécessaire pour déterminer les opérations les plus intéressantes à réaliser. Cette dernière analyse pour choisir la solution finale ne vise qu'un panel réduit de solution et peut directement faire appel au savoir expert.

3.2 Limitation – La solution est-elle l'optimisation ?

Nous avons vu qu'il existe des méthodes pour définir des solutions optimales, cette recherche est cependant rendue difficile par l'aspect variable des objectifs, des attentes du projet et de l'acteur considéré. On a tendance à recourir à l'optimisation, car l'étude devient complexe. Mais cela demande de poser le problème de façon claire et de déterminer toutes les propriétés que la solution doit respecter. Dans le cas du bâtiment, le problème est « ouvert », on souhaite optimiser et en même temps caractériser le bâtiment. On ne connaît pas ses capacités ni ce qui doit en être fait. L'utilisateur a alors le choix entre faire de fortes hypothèses sur ce que doit être la solution et avoir à choisir manuellement parmi un échantillon réduit.

La démarche est contraignante, elle demande de nombreuses évaluations du système sans pour autant apporter de connaissances sur celui-ci. Si les facteurs d'optimisation sont mal choisis, on obtient une solution qui ne correspond pas tout à fait aux attentes. Il n'est ensuite pas possible de vérifier le résultat du processus d'optimisation, car justement ce processus a été mis en place à cause de la complexité de l'analyse. Ainsi on ne sait pas comment et pourquoi l'optimisation est arrivée à cette solution. Il n'est également pas possible de prendre en compte le phasage des travaux, car ce sont des opérations partielles qui ne font pas partie du front de Pareto.

Comme nous l'avons vu en section 2.1, la grande majorité des opérations se fait avec un budget limité. L'optimisation multicritère peut effectivement intégrer une évaluation du coût de travaux en fonction de l'état initial et des modifications apportées, mais définir un critère financier peut mener à une autre solution. Une solution dont l'enveloppe financière disponible sera optimisée, mais qui ne correspond pas forcément à une étape de l'objectif final qui est visé sur le long terme. De plus, le phasage demande une analyse des aides que l'Etat fournit pour favoriser la réhabilitation et l'intégration des problématiques relatives à la mise en place des opérations successives.

De même, si une modification est faite après optimisation ou si la solution doit être adaptée lors de la mise en œuvre, le comportement de la solution est difficilement

connu et les conséquences sont à réévaluer. Il est tout de même possible dans certains cas de réaliser une étude de sensibilité avec les résultats issus de l'optimisation en considérant un léger biais (Marseguerra et al., 2003) – notamment dans une recherche par algorithme génétique où les populations se centreront autour de la solution optimale. Cependant, il est plus naturel de réaliser une analyse de sensibilité avant la recherche d'optimum. Cela permet de réaliser une première évaluation de l'importance relative des entrées et ainsi de sélectionner un jeu réduit de variables pour l'optimisation. Nous essaierons de développer cette idée au maximum pour savoir si analyser et comprendre le comportement du modèle peut nous permettre de choisir directement une solution.

« Dans la plupart des méthodes ayant recours à la pondération, le choix du décideur se limite à exprimer ses préférences entre les différents critères. Seul, le choix de la pondération peut alors être argumenté. Cela évacue complètement le véritable enjeu de la décision : le choix des actions » (Taillandier, 2009)

3.3 Vers plus de transparence

L'optimisation produit une étude incomplète, elle manque de transparence et ne prend pas en compte tous les facteurs. Dans tous les cas, le choix d'une solution passe nécessairement par un avis expert. De plus, les outils d'optimisation ne sont en réalité que très rarement utilisés dans les projets.

Pour mettre en place une autre démarche pour la recherche de solution les éléments de base de l'analyse ont été définis :

- Evaluer un panel important de solutions.
- Automatiser et rendre l'analyse aussi générique que possible, c'est-à-dire indépendante du bâtiment et du projet considéré.
- Prendre en compte l'état initial, l'amélioration faite sur les performances et les performances maximales atteignables.
- Donner les moyens d'argumenter les choix réalisés.
- Valoriser au maximum les simulations.

Il a ainsi été décidé d'optimiser la compréhension du modèle plus que la recherche de solution particulière. L'objectif de la démarche est de fournir des outils d'analyse qui permettent d'aider l'expert dans la recherche de solution et ainsi de mettre en place des mécanismes qui facilitent la compréhension du système bâtiment pour un gain en transparence.

Informé sur le gain de performance apporté par l'opération n'est pas suffisant. Il est nécessaire d'estimer le maximum de performance réalisable et de le comparer à l'état structurel du bâtiment afin de définir les objectifs à atteindre à plus ou moins long terme. La recherche doit permettre de trouver les voies d'amélioration qui nous mènent vers cet optimum tout en contrôlant les résultats intermédiaires.

La consommation annuelle n'est alors pas la variable la plus intéressante à analyser, c'est un résultat qui dépend de trop de paramètres. Pour l'analyse, il peut être intéressant de prendre le problème à l'envers, l'épaisseur de l'isolation est un élément sur lequel il sera possible de jouer et qui influe sur la performance du logement. Il est donc intéressant de déterminer quel est son rôle dans le bâtiment et dans quelle mesure il influe sur chaque sortie.

Les principales questions qui sont posées sont :

- Quel sera l'effet de la modification d'un paramètre d'entrée ?
- Pour modifier une sortie sur quel paramètre est-il nécessaire d'influer ?
- Quelles sont les performances maximales du bâtiment ?
- Quelle est la plage d'efficacité de chaque variable ?
- Que valent les résultats obtenus ?

Ainsi, nous ne voulons pas des ordinateurs pour concevoir à notre place, mais des ordinateurs pour nous aider à explorer des idées et d'être une aide à la conception intelligente. La méthodologie pour la recherche de solution qui est proposée se base sur l'utilisation des outils d'analyse de sensibilité pour étudier le comportement du bâtiment. L'analyse de sensibilité sur le bâtiment apporte des informations qui peuvent permettre d'aller au-delà de la recherche de solution. Ces méthodes fournissent un indice de confiance sur les résultats et aident également la calibration de modèle en déterminant les paramètres d'entrée les plus influents sur la sortie du modèle considérée (Mara et Tarantola, 2008).

Les logiciels de STD permettent de réaliser des estimations précises avant travaux mais il y a des différences entre prévision et réalité et peu de retours d'expérience sont disponibles sur ces écarts. Apporter de la fiabilité dans les estimations permettra de valider les solutions et de justifier leurs financements. Cela permet donc d'avoir une plus grande marge de manœuvre et aller plus loin dans la performance. Il faut s'assurer de la faisabilité des solutions étudiées et de leurs bonnes réalisations. Une étude stochastique doit être menée afin d'évaluer le risque de défaillance. Car chaque étape de la démarche apporte son jeu d'incertitudes : diagnostic, modèles, mise en

œuvre, occupation réelle, météo, *etc.* La question de l'analyse d'incertitude s'impose alors d'elle-même.

La confiance d'un résultat est déterminée par l'incertitude qui est associée, mais il est nécessaire pour cela de définir les incertitudes d'entrée. D'autre part même avec des entrées correctes il peut persister une erreur, qui est alors considérée comme une erreur systématique. C'est un biais du modèle qui peut correspondre à un phénomène physique qui n'a pas été pris en compte. L'évaluation de cette erreur n'est possible que si une évaluation de l'incertitude a déjà été réalisée.

4 Indicateurs de performance

La démarche de performance doit se baser sur des points d'évaluation, des variables d'évaluation qui représentent la capacité du bâtiment à répondre aux attentes. Cette évaluation est similaire à la démarche qualité utilisée dans les entreprises dans laquelle la création d'indicateur permet cette analyse.

4.1 Méthodologie de la création des indicateurs

Les normes ISO 9000 permettent de mettre en place un suivi de la qualité par la mesure de la performance. Cette performance peut prendre des formes très variées et inclure des définitions complexes. A partir de la définition de l'objectif global il est nécessaire d'identifier un ensemble d'objectifs élémentaires qui renseigneront de l'atteinte, plus ou moins partielle, de l'objectif global. Un système d'indicateurs appelé tableau de bord est alors mis en place pour fournir une ou plusieurs expressions de performance. Chacun des indicateurs comprend une mesure et un critère qui permet d'évaluer l'accomplissement de l'objectif élémentaire. La mesure de la performance globale est alors obtenue par l'agrégation de ce système d'indicateurs.

Nous avons vu en section 3 ci-dessus que la partie agrégation de ce système d'indicateur nécessite une forte connaissance du problème et peut être ainsi inadaptée à la recherche de solution. Nous considérons donc dans ce travail les différents indicateurs élémentaires sans produire d'évaluation globale.

La norme (FD X 50-171, 2000) indique les phases du cycle de vie des indicateurs :

- Identification
- Conception
- Mise en œuvre
- Exploitation
- Examen périodique

L'identification consiste à repérer au sein du processus, les *variables* relatives à l'atteinte de l'objectif. Les plus importantes sont identifiées pour former les objectifs élémentaires. Dans la phase de conception, chacune des variables essentielles produit une expression de performance. Toutes les contraintes sont retranscrites sous forme d'un ensemble correspondant de critères pour former les indicateurs. La Mise en œuvre et Exploitation correspondent réciproquement à l'application dans l'entreprise du tableau de bord ainsi créé et à la représentation et à l'utilisation des résultats. L'examen périodique permet d'évaluer l'évolution des processus et des performances, car ceux-ci peuvent rendre inadapté/désuet un indicateur ou faire apparaître un nouveau besoin. Il est alors nécessaire de revisiter les indicateurs précédemment décrits.

Il est important de maintenir le système d'indicateurs pertinent et cohérent. Dans une entreprise, les évolutions sont progressives et la révision du tableau de bord se fait sur de grandes périodes. Dans le cas présent, l'évolution est immédiate et nécessite donc une forte adaptabilité. Nous présenterons donc un ensemble non exhaustif de variables d'évaluation centrées sur les aspects thermiques uniquement.

Dans le cas général pour que les indicateurs soient de bonne qualité les normes précisent qu'ils doivent être construits afin de respecter certaines règles. Pour l'application de la méthodologie à un outil de simulation numérique, nous retiendrons qu'il doit :

- être fidèle et représentatif,
- mettre en évidence les évolutions,
- être facile à établir, alimenter et utiliser.

4.2 Les principaux points d'évaluation

Notre objectif premier est d'aider le concepteur à trouver quelles sont les capacités de réhabilitation du bâtiment étudié. Les indicateurs de sortie seront donc centrés sur les aspects thermiques. Pour une recherche de solution plus exigeante, il est évident que d'autres critères devront être pris en compte, mais ceux-ci ne seront pas abordés.

Les indicateurs de performance qui vont être sélectionnés doivent permettre de retranscrire le comportement du bâtiment. L'objectif est d'obtenir une représentation claire de multiples phénomènes temporels afin de permettre de différencier deux jeux de valeurs d'entrée ou de quantifier un phénomène.

Deux choix s'imposent:

- Utiliser des indicateurs classiques pour une évaluation globale de la performance, l'analyse de sensibilité permettant par la suite de lier le comportement des indicateurs en fonction des entrées.
- Utiliser des indicateurs plus explicites pour évaluer de façon plus précise l'effet d'un élément.

Les parties suivantes décrivent donc quelques indicateurs classiques ainsi que des indicateurs plus spécifiques.

4.2.1 *Evaluation générale – Impact des saisons définition des périodes d'étude*

Dans la pratique, des simulations annuelles au pas de temps horaire sont réalisées afin d'obtenir une évaluation globale du bâtiment. L'indicateur classique d'estimation des performances énergétiques est le besoin d'énergie en kWh/m².an. Cet indicateur est une estimation réglementaire annuelle qui permet de comparer tous les logements suivant leur climat et impose pour cela des valeurs normatives pour les scénarios et les paramètres de simulation (EN 15217, 2008). Cette sortie de la simulation thermique dynamique agrège de nombreuses données et offre ainsi une bonne évaluation du comportement global. Elle permet une comparaison avec les valeurs normatives de référence, c'est-à-dire les indicateurs de consommation classique usuellement employés par les professionnels.

Cependant, dans le cas présent l'indicateur est faussé par des hypothèses de simulation non réglementaires. Et la comparaison avec des valeurs réglementaires donnera des informations incorrectes. De plus, l'agrégation des données temporelles induit une perte d'information et il n'est donc pas possible de l'utiliser pour réaliser un dimensionnement.

Il est proposé d'évaluer cet indicateur sur de plus courtes périodes et de différencier le fonctionnement par saison. La Figure I-3 présente les évolutions annuelles classiques des températures intérieures et extérieures. On peut identifier trois grands modes de fonctionnement :

- Hiver: une quantité importante d'énergie est nécessaire afin de maintenir des conditions de température intérieures minimales.
- Eté: la température extérieure est assez élevée pour arrêter tous les systèmes de chauffage et la température intérieure est fortement liée à la température extérieure
- Intersaison la météo peut être très variable, les besoins sont alors intermittents et les deux scénarios précédent peuvent se présenter.

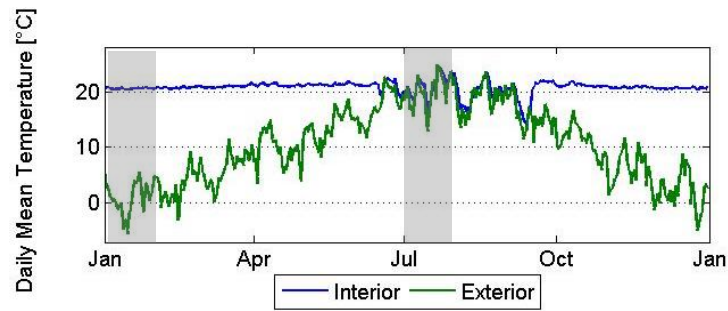


Figure I-3: Evolutions typique des températures intérieure et extérieure

Basé sur ces règles, le bâtiment sera évalué sur une période en l'été et en l'hiver afin de fournir une meilleure compréhension des comportements aux situations limites d'un fichier météo moyen. Le dimensionnement est ainsi réalisé en fonction de la variabilité du site sans prendre en compte un événement issu d'une météo extrême.

D'autre part, il a été nécessaire de considérer des périodes assez longues pour ne pas trop perturber les résultats. Lors du démarrage d'une simulation, le logiciel ne connaît pas la météo des jours précédents et ne sait pas dans quel état doit se trouver le bâtiment. L'initialisation permet de définir les températures initiales des zones et des parois. Si cette initialisation n'est pas faite correctement, le bâtiment démarre dans un état de déséquilibre et va mettre de nombreux jours à se stabiliser. L'initialisation permet de se rapprocher au maximum de cet état d'équilibre. Cependant, cet état de départ peut tout de même produire un léger biais, il est donc nécessaire de réaliser des simulations sur une période suffisamment longue pour pouvoir négliger cet effet.

Une période d'un mois en janvier et en juillet a semblé être un bon compromis. Cette durée dépasse largement la constante de temps du bâtiment, mais reste raisonnable pour supposer que les conditions climatiques sont stables et représentatives des variations de la saison.

La variable d'évaluation produite est donc la consommation énergétique en kWh/m².mois et ne concerne que les besoins de chauffage, refroidissement et les systèmes de ventilation. La prise en compte des autres dépenses énergétiques n'a pour l'instant pas d'intérêt, car elles correspondent à des valeurs forfaitaires. La valeur visée est la plus basse possible mais doit tout de même permettre d'assurer les conditions de confort.

4.2.2 *Evaluation du Confort*

Le confort est généralement associé au confort thermique, mais peut s'étendre à de nombreux domaines. La notion de confort est difficilement estimable. Celle-ci est généralement définie comme un état d'absence d'inconfort, mais la limite entre confort et inconfort est délicate à estimer. Il est ainsi possible d'identifier une situation d'inconfort, mais sans que cela ne produise pour autant une réaction de la part des occupants. Cela est dû aux aspects contextuels et sociologiques de l'occupation. Ces aspects s'observent particulièrement lors de l'analyse des conditions d'ouverture et de fermeture des ouvrants par les personnes présentes dans un local où il y a apparition puis disparition d'une situation d'inconfort thermique.

L'évaluation du confort est donc fortement incertaine sur des échantillons réduits de personne. Une évaluation classique du confort via la norme sur les critères d'ambiance (EN 15251, 2007) sera donc réalisée. Les points qui ont été étudiés touchent l'acoustique, la qualité de l'air, la thermique et le visuel. Les parties suivantes abordent tous ces points, mais seuls les indicateurs thermique et visuel sont utilisés par la suite.

A Acoustique

L'aspect acoustique a été jugé important, mais fait intervenir d'autres méthodes de résolution qui ne sont pour l'instant intégrées à aucun outil de conception thermique.

En réhabilitation, l'acoustique intervient dans les cas où une gêne est exprimée par les occupants et également dans le dimensionnement des éléments de façades. Cependant, certaines opérations conduisent à une dégradation du confort acoustique malgré une amélioration de l'isolation des bruits extérieurs par les vitrages. Cette augmentation de l'inconfort dépend entre autres de la source des bruits perçus. Un bruit de route même s'il est jugé gênant est généralement mieux accepté que les bruits d'occupation. Ainsi une amélioration de l'acoustique nécessite une étude approfondie fortement liée aux matériaux de la structure et aux cloisonnements. Ce type d'opération est alors plutôt destiné aux réhabilitations lourdes réalisées en site inoccupé.

En cas d'évaluation précise de celui-ci les paramètres importants du bruit sont, l'intensité, la durée et le spectre de la source. Les critères exprimés correspondent aux différents niveaux de bruit dans les zones sensibles comme les chambres et le séjour.

Tableau I-2 : Tableau d'évaluation normatif de l'intensité du bruit dans quelques locaux (EN 15251, 2007)

Type d'espace	Plage type [dBA]	Valeur de conception [dBA]
Salle de séjour	25 à 40	32
Chambre	20 à 35	26

B Qualité de l'air

La qualité de l'air est un élément qui est de plus en plus mis en avant. Elle joue un rôle sanitaire important afin de limiter les concentrations internes : COV, Radon, Odeurs, CO, Humidité, *etc.* Un contrôle de tous les polluants n'étant réalisable que si une évaluation précise des débits est faite, il est nécessaire de veiller à une très bonne étanchéité du bâtiment, de disposer de systèmes mécaniques à débit contrôlé et de vérifier la manipulation des ouvrants. C'est donc un point complexe à évaluer et depuis de nombreuses années des débits minimaux de ventilation sont imposées afin de pallier toutes ces problématiques. Il est alors généralement supposé que le respect de ces limites permet d'éviter tout problème majeur.

Les logements anciens réalisent généralement cette ventilation de façon non contrôlée. La modification de l'étanchéité joue alors sur la capacité du bâtiment à se ventiler naturellement. Il est important de maintenir les débits et au besoin ajouter une ventilation mécanique afin de répondre au manque de ventilation induit par la réduction de la perméabilité à l'air. En pratique, la seule mise en conformité est une tâche délicate qui n'est pas toujours possible. Et d'autre part, certaines problématiques sont dues aux occupants avec généralement obturation des bouches d'extraction ou des ouvertures au niveau des fenêtres. Un indicateur de ce type en réhabilitation ne serait pas exploitable.

Ce type d'intervention est donc à la discrétion du concepteur qui déterminera dans quelle mesure agir sur les débits en fonction du type de bâtiment et des problématiques présentes.

C Thermique

Evaluer la sensation de confort thermique est une tâche délicate. De nombreuses méthodes d'estimation ont été développées et comme le montre la Figure I-4 leur nombre ne cesse d'augmenter au fil des ans (Carlucci et Pagliano, 2012). Cela est dû à la définition très subjective du confort qui est le produit d'une interaction entre les domaines physiologique, physique, psychologique et sociologique et qui nous en donne ainsi une représentation propre (Moujalled, 2007). Le ressenti est différent

d'une personne à l'autre et il est difficile d'évaluer avec précision l'activité et l'habillement des personnes.

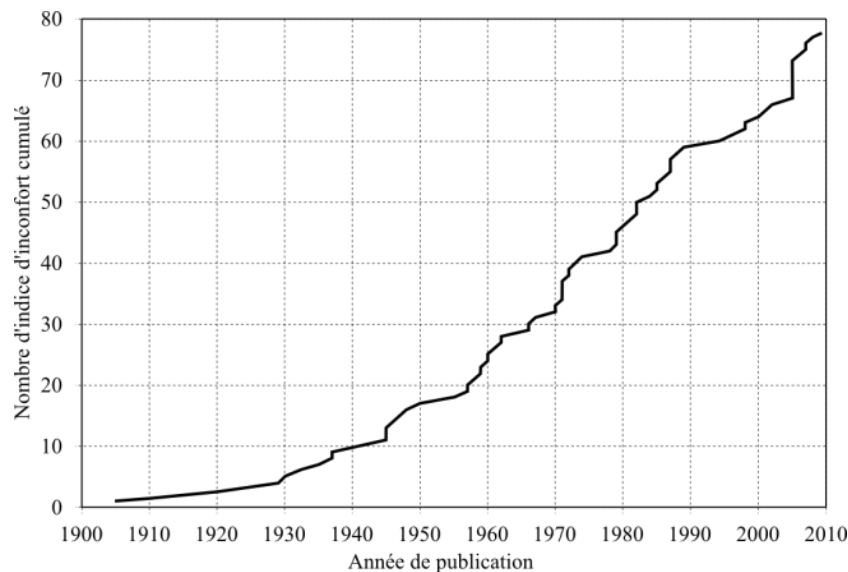


Figure I-4 : Evolution du nombre d'indices d'inconfort

Le nombre de facteurs pris en compte dans les méthodes est très variable. Certaines peuvent avoir recours à un nombre important de paramètres afin de retranscrire avec précision les phénomènes physiques, mais les méthodes généralement utilisées se limitent à quelques facteurs et font l'hypothèse que les conditions sont stables aussi bien pour l'ambiance que pour les occupants. Dans ces dernières, il est alors supposé que l'occupant est en équilibre avec son environnement, c'est-à-dire qu'il est dans l'ambiance depuis un certain temps et que son activité reste constante. On s'aide alors de questionnaires et de relevés sur d'importants échantillons de population et dans de nombreuses situations pour mettre en place des corrélations.

Cette évaluation du confort au niveau de l'ambiance peut être complétée par des conditions de confort plus locales. L'ISO 7730 (EN ISO 7730, 2006) précise des limites à respecter concernant les courants d'air, la dissymétrie de la température radiante, les gradients de température verticaux et les températures du plancher. Cette évaluation locale, peut permettre de détecter des situations d'inconfort qui ne sont pas identifiées dans les évaluations globales. Cependant, les logiciels font des évaluations zonales des ambiances et produire un indicateur de confort local ne semble donc pas pertinent.

Deux types d'indicateurs sont généralement utilisés pour évaluer le confort dans le logement. Ceux basés sur le PMV (Predicted Mean Vote), qui évaluent la sensation de confort dans une ambiance contrôlée où un certain nombre de paramètres doivent être estimés (habillement, métabolisme (activité), température, humidité, vitesse

d'air). D'autres, basés sur le confort adaptatif, qui considèrent que la température optimale du logement est principalement conditionnée par la température moyenne extérieure. L'occupant va s'adapter à son environnement et acceptera des conditions de confort plus larges.

Le PMV (Predicted Mean Vote) (Fanger en 1970) se base sur une évaluation de la sensation de confort obtenue dans des chambres climatisées. Le modèle a été obtenu à partir de corrélations issues du croisement des données de base des occupants et de l'ambiance avec les réponses sur leur état de confort. Pour une situation donnée, le confort est estimé entre -3 (très froid) et +3 (très chaud) où 0 est l'optimum (ni chaud, ni froid). Le PPD (Predicted Percentage Dissatisfied) est l'extension statistique du nombre de personnes insatisfaites variant de 5 % (PMV de 0) à 100 %. Ainsi même pour une situation idéale du point de vue du PMV il y aura toujours une part de personne en inconfort, ce qui représente bien l'aspect personnel.

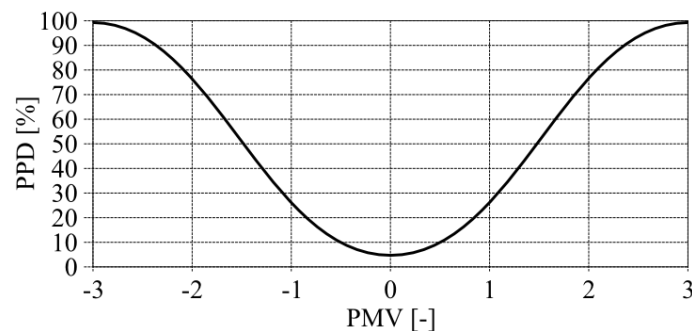


Figure I-5 : Relation PMV- PPD (Fanger)

L'indicateur normatif est donné par la catégorisation de cet état d'inconfort (EN ISO 7730, 2006). Les conditions intérieures sont estimées confortables si la catégorie II est respectée, soit moins de 10% d'insatisfaits :

- Cat I : $PMV < \pm 0,2$ soit $PPD < 6\%$
- Cat II : $PMV < \pm 0,5$ soit $PPD < 10\%$
- Cat III : $PMV < \pm 0,7$ soit $PPD < 15\%$
- Cat. IV : $PMV > \pm 0,7$ soit $PPD > 15\%$

Par définition, le PMV est plus adapté à la saison de chauffe. Car les paramètres intérieurs sont mieux contrôlés. Les fenêtres sont fermées la plupart du temps, le débit de ventilation est donc au minimal admissible –modulo les débits d'infiltration– et la température sera maintenue à une valeur définie par la régulation du système de chauffage. Reste les données relatives à l'occupant qu'il est nécessaire d'estimer.

Malgré le fait que le PMV fasse partie des méthodes généralement employées, celui-ci ne sera pas utilisé, car il demande une évaluation du métabolisme et de

l'habillement. Le choix de ses facteurs est une source d'erreur importante qui risque donc de fausser les résultats. D'autre part, l'évaluation est faite en PMV et cette donnée est difficilement interprétable, car c'est une évaluation qualitative de l'état de confort.

Le confort adaptatif (Nicol and Humphreys en 1972) est donc préféré. Il se base sur l'hypothèse que la condition de confort dépend grandement du contexte, du comportement des occupants et de leurs attentes. Les occupants n'ont pas besoin de conditions fixes, mais s'adapteront (habillement, posture, métabolisme, activité, ...) aux variations climatiques s'ils en ont les moyens avec notamment un accès libre aux ouvrants. Dans un logement réel et naturellement ventilé, les conditions de confort sont alors plus larges et dépendent principalement de la température extérieure (Nicol et Humphreys, 2002). Il a été montré qu'une température optimale de confort –opérative- pouvait être estimée par une relation linéaire avec la température moyenne extérieure (Figure I-6).

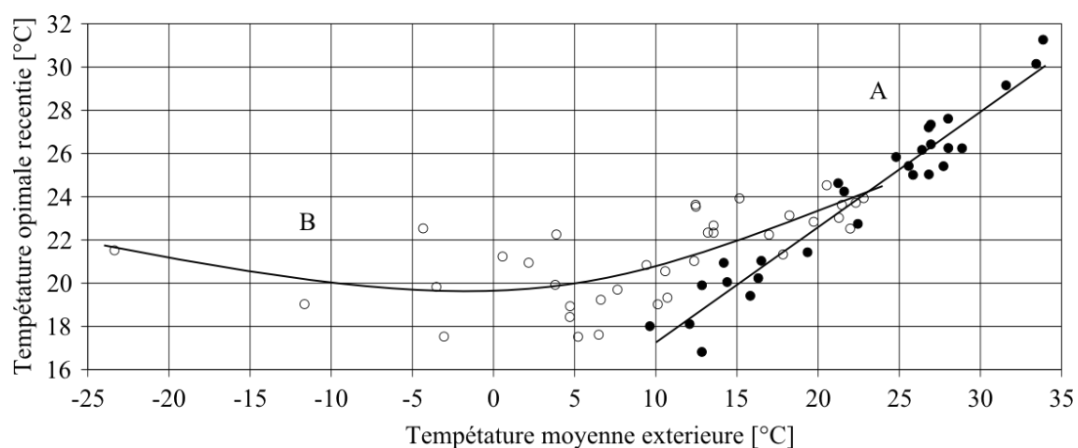


Figure I-6 : Relation entre la température optimale et la température extérieure (Nicol and Humphreys) (A) logements en évolution libre. (B) Logements conditionnés par un système

Cette définition du confort est utilisée réglementairement pour évaluer uniquement le confort d'été, car une très bonne corrélation est obtenue pour les bâtiments en évolution libre. Mais initialement, celui-ci s'entend sur une plus grande plage de température extérieure et reste un indicateur viable pour les situations hivernales (Nicol et Humphreys, 2002). Cet indicateur permet d'obtenir une représentation claire de l'état de confort par un lien direct avec la température intérieure.

L'indicateur normatif est donné par la catégorisation de l'écart entre la température intérieure et la température de confort. Les conditions intérieures sont estimées confortables si la catégorie II est respectée. Les pourcentages de satisfaction ont été définis lors des discussions pour la mise en place de la norme EN15251(Boerstra,

2006; DRAFT prEN 15251, 2005) à partir des travaux précédant (Dear et Brager, 1998; de Dear et al., 1997).

- Cat I : $\Delta T < 2^{\circ}\text{C}$ (90% confortable)
- Cat II : $\Delta T < 3^{\circ}\text{C}$ (80% confortable)
- Cat III : $\Delta T < 4^{\circ}\text{C}$ (65% confortable)

La méthode donne des informations horaires sur la sensation de confort pour chaque ambiance et il est nécessaire d'agréger les résultats afin d'obtenir une estimation simple de l'inconfort général sur l'ensemble de la période. Différentes méthodes peuvent être employées pour obtenir une évaluation scalaire du confort. Classiquement, des limites sont définies pour différentes catégories de confort plus ou moins contraignantes. Une catégorie est validée, si durant la période d'étude ses limites ne sont pas dépassées plus de 3 à 5% du temps d'occupation sur au moins 80 à 95% de l'espace. Un dépassement occasionnel du critère d'inconfort est ainsi permis, mais on ne sait pas dans quelle mesure la limite est respectée ou dépassée.

Afin de prendre en compte ce dépassement, nous avons choisi la théorie de Sicurella (Sicurella et al., 2011). Dans un premier temps, l'évolution de la température moyenne opérative est calculée en fonction du volume des zones chauffées. Ensuite, la qualité de l'ambiance est évaluée via l'ITD (Intensity of Thermal Discomfort) qui représente l'intégrale de la différence entre la température opérative évaluée et la limite de confort définie. Cet indicateur fournit des degrés jours au-delà et en deçà des limites pour chaque période considérée. La Figure I-7 présente l'évolution de ces écarts pour une journée, mais le principe peut être appliqué sur de grandes périodes d'étude. La zone de confort de Classe I est définie comme limite, elle correspond à l'état idéal avec 90% des personnes satisfaites et toute variation entre ces limites n'apporte pas d'information. Le dépassement est autorisé et évalué afin de quantifier les variations autour de la Classe II, définie comme objectif de confort. Aucun indicateur sur la fréquence des dépassements n'est évalué, car on considère les conditions relativement homogènes sur les périodes d'étude.

$$\begin{aligned} \text{ITD}_{\text{over}} &= \int_P \Delta T_{\text{over}}(\tau) . d\tau \\ \Delta T_{\text{over}}(\tau) &= \begin{cases} T_{\text{op}}(\tau) - T_{\text{over}} & \text{if } T_{\text{op}}(\tau) > T_{\text{over}} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (\text{I.2})$$

$$ITD_{\text{under}} = \int_P \Delta T_{\text{under}}(\tau) . d\tau$$

$$\Delta T_{\text{under}}(\tau) = \begin{cases} T_{\text{under}} - T_{\text{op}}(\tau) & \text{if } T_{\text{op}}(\tau) < T_{\text{under}} \\ 0 & \text{otherwise} \end{cases} \quad (I.3)$$

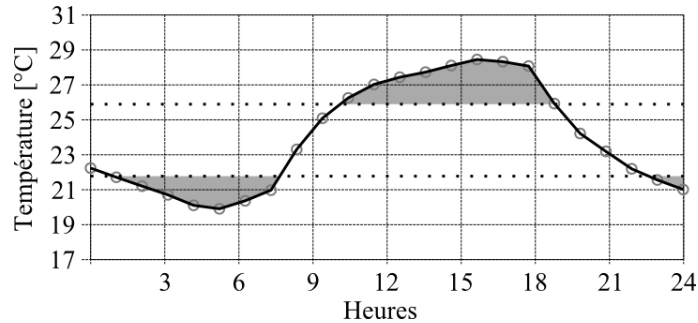


Figure I-7 : Définition de l'intensité de l'inconfort – ITD (Sicurella)

D Visuel

Les conditions visuelles sont généralement étudiées dans le tertiaire où l'éclairage peut être un poste de consommation important et où il est nécessaire de répondre à des besoins d'éclairement particuliers et très bien définis. Ainsi de nombreux travaux ont été réalisés en éclairagisme et particulièrement sur l'étude des besoins en éclairage ainsi que des sources d'inconfort au niveau des postes de travail. L'éclairage naturel a été étudié plus récemment afin de limiter au mieux la part d'éclairage artificiel nécessaire au respect des conditions de travail. L'EN 15251 précise que pour des raisons de santé, de confort et d'énergie il est préférable d'utiliser la lumière du jour avec éventuellement un éclairage d'appoint plutôt que la lumière artificielle.

De façon générale, la littérature présente de nombreux critères permettant d'évaluer la qualité visuelle d'une ambiance par l'intensité, la couleur, l'uniformité, l'éblouissement, l'influence de l'éclairage naturel, etc. Toutes ses données sont évaluées afin que l'occupant puisse réaliser la tâche définie sur son poste durant ses horaires de travail en limitant la fatigue visuelle et les risques d'erreur.

Le cas du logement apporte quelques spécificités, il n'y a pas de postes ou de surface à optimiser pour une tâche donnée, l'occupant adaptera son espace en fonction de ses besoins, de son goût, de ses envies. Les périodes d'occupation sont très variables et principalement en fin de journée quand les apports solaires diminuent. Les conditions de confort dans un logement ne sont donc pas les mêmes que dans un espace de travail et rien n'est réglementairement imposé pour l'habitation. De plus, dans la pratique, les luminaires dans les logements sont peu ou pas optimisés, car ils occupent également un rôle décoratif. Cela renforce l'importance de l'éclairage

naturel, car il devient alors une source de confort non négligeable. Cependant, si un éclairage trop faible peut être corrigé par l'utilisation des luminaires, trop de rayonnement solaire induit un éblouissement en intersaison et des surchauffes estivales et ses conséquences ne sont pas toujours corrigibles.

Dans cette partie, nous décrirons exclusivement la partie naturelle de l'éclairage. L'éclairage artificiel sera laissé à l'initiative des occupants, car les points d'éclairage sont généralement déterminés et non modifiables et il y a trop d'inconnues pour pouvoir proposer des solutions d'éclairage. Ce complément artificiel ne pourra être traité que de façon forfaitaire pour les apports thermiques. La mise en place de l'évaluation doit permettre de caractériser l'ambiance lumineuse naturelle et observer l'impact d'une modification du type de vitrage ou des protections solaires.

L'idéal au niveau du confort serait d'utiliser la luminance de fond (lumière diffusée par les surfaces intérieures en $[cd/m^2]$) présente lors de l'occupation en début/fin de journée. C'est essentiellement ce paramètre qui nous fait dire qu'un local est lumineux ou non. Malheureusement, ce n'est pas un indicateur classique, donc les travaux sur son influence sont rares et les valeurs de référence pour juger de l'acceptabilité manquent. Cette valeur est également fortement dépendante du revêtement des parois ce qui la rend incertaine.

Une évaluation plus classique sera donc utilisée afin de disposer d'un retour d'expérience plus important, mais il sera tout de même nécessaire d'adapter des indicateurs. Dans la littérature, l'éclairage –quantité de lumière arrivant sur une surface plan située à une hauteur de 80 cm en $[lux]$ – est évalué durant les périodes de travail, pour générer une carte pour chaque pas de temps de sa simulation. Cette carte est utilisée comme base pour caractériser le local, mais elle permet également d'estimer un risque d'éblouissement quand les valeurs dépassent une certaine limite. A partir de cette carte, de nombreuses informations peuvent être obtenues sur l'ensoleillement : moyenne, nombre d'heures, carte des maximums, carte moyenne d'ensoleillement, *etc.*

Le Facteur de lumière du jour (Daylight factor) est le plus courant, il donne le rapport entre éclairements intérieurs et l'éclairage extérieur sur des plans horizontaux et avec des conditions de ciel particulières. Il permet ainsi d'évaluer la qualité de la conception, mais ne fournit pas une évaluation sur site. L'Autonomie de lumière du jour (Daylight Illuminance) est une autre mesure courante qui correspond au pourcentage de temps d'occupation où l'éclairage en tout point est supérieur à un minimum.

Pour l'évaluation, nous reprenons la définition de l'Autonomie de lumière du jour utile (Useful Daylight Illuminance) introduite par Nabil et Mardaljevic (Nabil et Mardaljevic, 2006) qui ajoutent une limite haute pour l'inconfort d'éblouissement. Pour les valeurs minimales et maximales d'éclairement beaucoup de sondages déterminent les limites entre 100 et 2000lx. Mais il a également été observé que les personnes ont tendance à tolérer des niveaux d'éclairement beaucoup plus faibles lorsqu'il s'agit de la lumière du jour. Ce phénomène est particulièrement visible en fin de journée quand les occupants continuent de lire à des niveaux de clarté aussi faibles que 50 lux. Il est donc très difficile de donner une limite à respecter quand il n'y a pas de tâche définie. Il est également nécessaire de prendre en compte le modèle d'évaluation, la simulation donne des valeurs horaires et en fin de journée l'éclairement diminue très rapidement. Il n'y a donc qu'une différence fixe de deux heures par jour –une heure matin et une heure le soir– entre une limite de 100 ou de 0 lux ce qui n'apporte pas vraiment d'informations. Les limites initiales qui déterminent le confort en 100 et 2000 lux sont donc utilisées.

La base de calcul retenu correspond à l'éclairement moyen de la zone, car aucun poste particulier de travail ne peut être défini, et sans les limites sur les horaires, car l'occupation est généralement concomitante avec les crépuscules en début et fin de journée où les conditions de confort ne peuvent naturellement pas être assurées. L'ambiance sera évaluée par un taux de confort égal au rapport entre le nombre d'heures de confort –avec ou sans occupation– sur le nombre d'heures d'ensoleillement, ainsi que par l'éclairement moyen sur la durée. Pour les surchauffes, une étude au niveau des bilans énergétiques des vitrages –détaillée dans la partie suivante- permettra de quantifier l'énergie solaire entrante de façon plus précise.

4.2.3 Détails de performance énergétique

Pour détailler un peu plus le fonctionnement du bâtiment, il est possible d'observer les principaux postes de dépense via les différents modes de transferts : Aéraulique pour les échanges avec l'air extérieur ; Conduction pour les échanges avec les parois ; Rayonnement pour notamment les apports solaires. Le rôle du bâtiment est ensuite de maintenir le volume d'air intérieur dans des conditions acceptables par l'intermédiaire des systèmes énergétiques. La simulation énergétique suit la même démarche et calcule les quantités d'énergie apportées et extraites de ce volume. Ainsi, si des conditions sont stables la relation suivante peut être considérée :

$$(\text{Aéraulique}) + (\text{Conduction}) + (\text{Rayonnement}) = (\text{Système}) + (\text{Charges Internes})$$

Ce sont les principaux modes de transfert, une analyse horaire de ceux-ci doit permettre de déterminer les principales voies de déperdition ou d'apport de chaleur.

Une information condensée de chacun de ces transferts est obtenue en réalisant une somme séparée des parties positives et négatives des flux. L'indicateur choisi est alors composé de ces deux sommes et permet d'évaluer les apports et pertes de chaque mode de transfert durant le mois d'étude, il est donc exprimé en kWh/mois.

Il est ensuite possible de détailler d'avantage. L'enveloppe est considérée comme l'ensemble des surfaces qui protègent la ou les zones chauffées sans prendre en compte les surfaces entre zones. L'énergie qui entre et qui sort de cette enveloppe est quantifiée via le flux convectif intérieur. Le choix de ce flux permet de s'affranchir des phénomènes externes –rayonnement solaire, météo– et internes –rayonnement intérieur. Ainsi les pertes dues à la surface en question pourront être comparées quel que soient la localisation ou le type de paroi mis en place. Le flux convectif intérieur est détaillée si besoin par type de surface et par orientation afin de comparer les pertes et apports entre d'une part les différentes faces du bâtiment et d'autre part entre les différents éléments et ainsi identifier les points faibles de l'enveloppe. Ces différentes valeurs sont exprimées en fonction de la surface de transfert considérée en kWh/m².mois. Les vitrages disposant en plus d'une part de rayonnement ont une seconde définition qui considère le flux total intérieur. Ce choix permet de mettre en balance la moins bonne résistance thermique des vitrages par les apports solaires en journée.

Ces indicateurs ne disposent pas de références ou de limites. C'est la comparaison des résultats qui permettra de faire une analyse.

5 Un besoin de méthode innovante pour la recherche de solution en réhabilitation

Cette partie a permis de mettre en avant que le parc de logement actuel est composé en grande partie de logements anciens et que les réglementations thermiques ont subi une évolution importante. Cette évolution nécessite de modifier les logements existants pour intégrer les problématiques énergétiques actuelles et les préparer aux contraintes futures. On peut alors s'interroger sur le potentiel d'augmentation des performances compte tenu du bâtiment existant étudié et de toutes les contraintes du projet. Les critères de performance peuvent être nombreux –différents types de confort, financier, consommation, énergie grise, *etc*– et leurs choix ainsi que l'importance associée à chacun dépend du niveau d'exigence du projet.

Trouver une solution optimale dans un domaine aussi subjectif que le bâtiment nécessite des méthodes innovantes, car l'arbitrage est difficile. La conception des bâtiments s'est toujours basée sur l'analyse et le retour d'expérience du concepteur. L'idée est de continuer dans cette voie et d'apporter au concepteur les outils d'analyse nécessaires pour évaluer le modèle thermique de bâtiment et faire des choix. L'objectif est de quantifier un potentiel et d'identifier les voies d'amélioration pour obtenir des performances élevées vis-à-vis des capacités du bâtiment.

Les simulations sont de plus en plus précises et les réalisations de plus en plus performantes. Et par conséquent, les attentes face aux résultats de simulation sont de plus en plus importantes. Il apparaît un besoin d'analyse et de quantification des erreurs entre la phase conception et la réalisation finale. Les techniques d'analyse de sensibilité permettent d'explorer le comportement du modèle et apportent une nouvelle piste pour la recherche et la validation de solution. Les ordinateurs actuels disposent de puissance de calcul suffisante pour nous permettre en une nuit de calcul de réaliser suffisamment de simulations pour évaluer le comportement du bâtiment sur, par exemple, l'ensemble des solutions de réhabilitation. L'objectif est de réaliser une étude purement thermique afin d'évaluer la recherche de solution tout en laissant la possibilité de prendre en compte les autres contraintes. La méthodologie s'axe autour de la modularité de l'outil en fonction des attentes et du projet considéré.

Le Chapitre II présente tout d'abord les techniques d'analyse qui seront choisies et utilisées dans la mise en place de cette démarche.

Equation Chapter (Next) Section 1

Chapitre II : OUTILS POUR L'ANALYSE DU COMPORTEMENT DES MODELES THERMIQUES

1 L'analyse de sensibilité pour le bâtiment

1.1 Les objectifs de l'analyse de sensibilité

Le domaine de l'analyse de sensibilité étudie les modèles pour estimer l'influence des entrées sur les sorties. Une large variété de méthodes a été développée afin d'identifier les variables qui ont un rôle prépondérant au sein des modèles en fonction des sorties étudiées et de quantifier ces influences. Dans la pratique, l'analyse de sensibilité est généralement couplée à l'analyse d'incertitude. Celle-ci permet à partir de données statistiques sur les variations d'entrée de correctement propager ces informations à travers le modèle afin d'obtenir les variations résultantes sur la ou les sorties. Si les variations appliquées sur les entrées sont la représentation de la méconnaissance du problème posé, les variations produites sur les sorties correspondent alors à l'incertitude produite par cette méconnaissance. Une vision commune de ces deux approches permet de quantifier l'influence des entrées par rapport à la méconnaissance présente en sortie. L'influence ainsi calculée fournit une représentation plus cohérente du rôle des entrées vis-à-vis d'une recherche de la fiabilité. Une entrée influente, mais connue précisément pourra alors disposer de la même importance vis-à-vis de l'erreur qu'une entrée peu influente, mais plus difficilement maîtrisable.

Les techniques d'analyse de sensibilité permettent ainsi de se diriger vers la fiabilité des résultats et peuvent aider à minimiser les écarts entre modèle et réalité. Elles sont également particulièrement adaptées pour apporter une compréhension du comportement du modèle par le biais d'informations qualitatives : effet linéaire/non linéaire, effet conjoint de plusieurs variables – interaction ; et ainsi de savoir comment les différentes parties du modèle interagissent (Saltelli et al., 1999). Elles

disposent de bases mathématiques solides et d'applications dans de multiples domaines tels que le nucléaire (Helton, 1993), l'hydrologie (Castaings, 2007) ou l'industrie (Rocquigny et al., 2008).

La thermique du bâtiment porte un intérêt croissant à l'application de ces méthodes. Elles sont utilisées afin d'analyser les modèles et permettent une identification préalable des variables les plus influentes afin d'accélérer la recherche de solutions par l'optimisation. De nombreuses études dans le secteur ont déjà été réalisées (Tian, 2013). Ces méthodes restent cependant limitées au domaine de la recherche et les logiciels de simulation thermique ne disposent pas encore d'options pour réaliser ce type d'étude.

Cependant, l'évolution des réglementations thermiques contraint le concepteur à analyser l'effet de toutes les variables et à prévoir les consommations en amont du projet et de façon de plus en plus précise. L'évolution des solutions techniques employées dans le bâtiment est généralement faite de façon progressive en se basant sur le retour d'expérience, mais les outils de simulation servent de plus en plus pour la recherche et l'évaluation préalable de solution. Il est donc certain que les techniques d'analyse continueront de se développer et de s'implémenter dans les outils de simulation pour la conception de bâtiment basse consommation.

Nous verrons dans ce chapitre quelles sont les méthodes de l'analyse de sensibilité et les évolutions récentes qui peuvent apporter une meilleure compréhension du bâtiment aussi bien pour la fiabilité des résultats que pour la recherche de solution.

1.2 Choix du type de méthode d'analyse de sensibilité pour l'étude des bâtiments

Le domaine de l'analyse de sensibilité englobe de nombreuses méthodes qui peuvent se regrouper en quelques grandes familles selon la complexité/régularité du modèle étudié et le nombre de simulations nécessaires (Iooss, 2011). Certaines familles de méthode ne seront donc pas adaptées ou simplement incompatibles en fonction du cas d'application. De plus, il est nécessaire de prendre en compte les besoins de l'étude afin de choisir la technique qui permettra d'extraire les informations voulues avec le plus faible coût de calcul.

Dans un premier temps, il est possible de discerner deux grandes catégories en fonction du domaine d'étude. Les premières sont les analyses de sensibilité locales dans lesquelles la réponse est étudiée autour d'un point de fonctionnement. Le comportement du modèle est supposé linéaire et le coût de calcul est faible. Ces

techniques sont généralement employées si le modèle permet un accès direct aux dérivées partielles des sorties. Dans le cas contraire, celles-ci sont évaluées par une variation des entrées avec une faible dispersion –1%–, cette approche n'est alors pas compatible avec des entrées trop incertaines, car les évaluations sont éloignées du point de fonctionnement et la linéarité locale n'est plus assurée. La deuxième catégorie correspond par opposition à toutes les techniques qui peuvent être appliquées sur la totalité du modèle et qui se regroupent sous le nom de méthodes globales.

Etant donné, que le problème porte sur l'étude des solutions de réhabilitation par un modèle de STD, les paramètres d'entrée subiront de grandes variations relatives à l'écart pouvant être présent entre un bâtiment ancien et un bâtiment récent. Nous utiliserons donc exclusivement des méthodes globales.

Dans un second temps, nous allons écarter les méthodes dites intrusives. Elles utilisent les équations du modèle comme sujet d'étude afin d'en extraire les informations de façon particulièrement efficace. Dans le cas de la simulation thermique, bien que tous les modèles soient connus et déterministes ceux-ci interagissent au sein d'un code de calcul indépendant et plus ou moins transparent en fonction de l'outil de modélisation. Face à cette complexité, le modèle est considéré comme une boîte noire, des données d'entrée sont fournies au modèle qui détermine alors les résultats de sorties associés. Les variables internes du modèle et les étapes du calcul ne sont pas utilisables, seule l'interprétation des variations de sortie par rapport aux variations d'entrée permet d'étudier le comportement. Il n'est donc pas possible de produire une expression analytique du problème ni des dérivées qui le composent. Nous écarterons donc les méthodes intrusives.

De nombreuses autres méthodes entrent dans les critères retenus jusqu'à présent, les plus connues sont les méthodes traditionnelles regroupées sous l'appellation générique méthode de Monte-Carlo. Une exploration aléatoire du modèle est réalisée et certaines informations sont extraites par une étude directe des résultats. Pour ne citer que les plus connues :

SRC Standardised Regression Coefficients, CC Correlation Coefficients, PCC Partial Correlation Coefficients ;

SRRC Standardised Rank Regression Coefficients, RCC Rank Correlation Coefficients et PRCC Partial Rank Correlation Coefficients.

Une rapide comparaison de l'ensemble de ces techniques est regroupée par Helton et Davis (Helton et Davis, 2002). Ces méthodes ne seront cependant pas abordées ici, car elles disposent de capacités inférieures aux méthodes présentées dans la suite (Sallaberry et Helton, 2006). L'analyse des effets est de moins bonne qualité dès que le comportement de la sortie devient complexe. Ces méthodes se basent sur des régressions ou des corrélations, l'efficacité est alors très dépendante de l'erreur R^2 – présentée dans le Chapitre IV :2.4– entre le modèle et le comportement estimé (Saltelli et al., 1993). Plus cette erreur est importante plus il y a une part de variation non expliquée pour chaque effet, cette variation résiduelle peut donc influencer sur l'estimation de l'effet des entrées. Or, étant donné que les sorties à observer ne sont pas toutes définies, il ne sera pas possible de faire d'hypothèses sur la linéarité ou la monotonie du comportement afin de s'assurer de la bonne utilisation de ces techniques. Il a ainsi été montré qu'il est préférable pour les analyses automatiques d'employer des méthodes basées sur l'analyse de la variance, car elles offrent une grande souplesse d'utilisation (Archer et al., 1997). De plus, celles-ci sont devenues très performantes, notamment sur le coût de calculs qui étaient leur frein principal grâce aux importantes améliorations qu'elles ont subies ces dernières années. Ces types de méthodes se rassemblent sous l'appellation ANOVA pour ANalysis Of VAriance et sont décrites dans la section suivante.

2 Théorie des techniques d'analyse employées

2.1 Description de la décomposition de la variance – ANOVA

Les méthodes basées sur la décomposition de la variance font partie de la famille des méthodes HDMR (High-Dimensional Model Representation). Dans ce type de représentation, la fonction est décomposée en une suite où chaque terme est le reflet d'une contribution indépendante ou coopérative des variables d'entrée sur la ou les sorties. Les méthodes ANOVA (ANalysis Of VAriance) analysent tout ou partie de ces termes pour en extraire les contributions des entrées étudiées. Nous décrirons dans cette partie la décomposition particulière utilisée par ce type de méthode.

Soit \mathcal{M} le modèle de simulation thermique dynamique, une fonction déterministe de \mathfrak{R}^m dans \mathfrak{R}^k avec m le nombre de variables et k le nombre de sortie :

$$\begin{aligned} \mathbf{y} &= \mathcal{M}(\mathbf{x}) \\ \mathbf{x} &= (x_1, x_2, \dots, x_m) \in \mathfrak{R}^m \\ \mathbf{y} &= (y_1, y_2, \dots, y_k) \in \mathfrak{R}^k \end{aligned} \tag{II.1}$$

Les variables sont les propriétés du bâtiment qui seront affectées par l'opération de réhabilitation telles que l'isolation, les vitrages ou le renouvellement d'air. Elles sont définies sur un domaine d'étude représenté par D_x . Les autres entrées du modèle non impactées par l'opération de réhabilitation éventuelle sont décrites comme des paramètres et sont maintenues à des valeurs fixes durant toute l'étude. Les informations relatives à la variation des entrées sont décrites sous forme de densité de probabilité. Chaque variable x_i dispose d'une fonction de répartition associée p_{x_i} qui informe sur la distribution des points qui seront analysés. On définit alors la densité du vecteur d'entrée \mathbf{x} comme le produit des densités marginales de chaque variable aléatoire :

$$p_{\mathbf{x}}(\mathbf{x}) = p_{x_1}(x_1) \cdot p_{x_2}(x_2) \cdot \dots \cdot p_{x_m}(x_m) \quad (\text{II.2})$$

On considère ensuite comme sortie, tout résultat de la simulation qui peut être défini sous forme de scalaire. Ces résultats correspondent aux indicateurs choisis en Chapitre I :4.2. Chacune de ces sorties est construite de façon indépendante et il est possible d'étudier une sortie à la fois sans que cela perturbe le fonctionnement du modèle ou l'étude qui est faite.

$$l \in [1; k], \quad y_l = \mathcal{M}_l(x_1, x_2, \dots, x_m) \quad (\text{II.3})$$

Certains auteurs se sont intéressés aux corrélations que peuvent avoir ces sorties et donc à la nécessité de reproduire les mêmes études sur des sorties pouvant être des combinaisons linéaires. La solution généralement retenue est une identification des modes propres des sorties afin de réduire la dimension de l'espace des résultats (Garcia-Cabrejo et Valocchi, 2014). L'analyse est ensuite réalisée de la même manière, mais sur chacun de ces modes propres. L'avantage de la non-redondance de l'étude se justifie d'un point de vue numérique, mais fait perdre le sens physique attribué à chaque sortie. D'autre part, pour l'application actuelle le temps de traitement est très faible par rapport au temps de simulation ce qui ne justifie pas non plus la mise en place d'une telle méthode. Il est donc préférable pour l'interprétation des résultats de garder des indicateurs clairs et construits pour avoir un sens physique déterminé.

Dans la suite, le modèle est alors supposé être une fonction de \mathfrak{R}^m dans \mathfrak{R} , les indices relatifs aux sorties sont donc supprimés et l'étude est répétée individuellement pour chaque indicateur.

$$Y = \mathcal{M}(x_1, x_2, \dots, x_m) \quad (\text{II.4})$$

L'analyse ANOVA provient de la propriété montrée par I.M. Sobol (Sobol', 1990) qui implique que toute fonction mathématique à carré intégrable peut se décomposer comme une somme de fonctions orthogonales. Le modèle se décompose donc de façon unique en une suite de fonctions de dimension croissante :

$$\mathcal{M}(x_1, x_2, \dots, x_m) = f_0 + \sum_{i=1}^m f_i(x_i) + \sum_{i < j}^m f_{ij}(x_i, x_j) + \dots + f_{1\dots m}(x_1, \dots, x_m) \quad (\text{II.5})$$

Le premier terme f_0 correspond à la valeur moyenne, puis une somme de fonctions élémentaires f_i qui dépendent de chaque variable d'entrée. Il s'ajoute ensuite les fonctions d'interactions d'ordre 2, 3, *etc.* qui font intervenir les différentes combinaisons de variables jusqu'à interaction d'ordre maximal qui correspond à la fonction qui fait intervenir toutes les variables. Le nombre de termes qui compose chacune de ces sommes est défini par le nombre de combinaisons possibles entre variables et peut être exprimé par les coefficients binomiaux. Le modèle est ainsi décomposé au total en 2^m fonctions.

On remarque à ce niveau qu'il est primordial de disposer d'entrées non corrélées afin de permettre la distinction de chacune de ces fonctions. Au niveau des équations, si deux variables x_a et x_b ont des variations corrélées, il ne sera pas possible d'identifier correctement les fonctions $f_a(x_a)$, $f_b(x_b)$ et leurs interactions $f_{a\dots}(x_a, x_{\dots})$ et $f_{b\dots}(x_b, x_{\dots})$. Les variables d'entrée choisies doivent pouvoir être fixées librement et être toutes indépendantes dans leurs variations. Dans le cas de la simulation thermique dynamique, des corrélations entre variables se retrouvent dans le fichier météorologique (température, humidité, rayonnements) ou dans les scénarios d'occupation (nombre de personnes, utilisation des équipements, ventilation hygro-réglable). Ainsi dans chacun de ces groupes, il n'est pas possible d'obtenir des variations indépendantes, à moins de perdre le sens physique donné aux valeurs des entrées. Des travaux pour permettre l'évaluation de la sensibilité de chaque entrée en considérant leur corrélation ont été menés récemment par G. Chastaing et S. Kucherenko (Chastaing, 2013; Kucherenko et al., 2012), la théorie reste cependant limitée à un nombre restreint d'entrées et l'interprétation des résultats obtenus s'avère délicate. Une alternative généralement employée consiste à considérer le groupe de variable comme une unique entrée tel que décrit par Saltelli et al. dans (Saltelli et al., 2007) paragraphe 1.2.15.

Le choix des entrées peut également permettre d'éviter une corrélation évidente. Par exemple, la conductivité thermique d'un matériau n'aura d'effet que conjointement avec l'épaisseur de celui-ci. L'effet réel de l'élément isolant est donc partagé entre ces

deux variables et sera plus difficile à évaluer si ces deux entrées sont étudiées conjointement.

Pour décrire plus en détail les fonctions exprimées en (II.5) vis-à-vis du modèle initial, le formalisme statistique sera employé. On définit ainsi un résultat conditionné, l'espérance et la variance tels que :

$$\begin{aligned} Y|\mathbf{x} &= \mathcal{M}(\mathbf{x}) \\ E(Y) &= \int_{D\mathbf{x}} \mathcal{M}(\mathbf{x}) p_{\mathbf{x}} d\mathbf{x} \\ V(Y) &= \int_{D\mathbf{x}} (\mathcal{M}(\mathbf{x}) - f_0)^2 p_{\mathbf{x}} d\mathbf{x} \end{aligned} \quad (\text{II.6})$$

Il est ainsi possible de définir $E(Y|x_i)$, l'espérance de Y conditionnée par une valeur x_i uniquement. Et avec cela les différentes contributions :

$$\begin{aligned} f_0 &= E(Y) \\ f_i(x_i) &= E(Y|x_i) - E(Y) \\ f_{ij}(x_i, x_j) &= E(Y|x_i, x_j) - E(Y|x_i) - E(Y|x_j) + E(Y) \end{aligned} \quad (\text{II.7})$$

Dans un cadre stochastique, chacun de ces termes explique donc une partie de la réponse du modèle et produit ainsi une variance partielle. La variance totale de la sortie peut alors être vue comme la somme de toutes ces variances :

$$V = \sum_{i=1}^m V_i + \sum_{i<j}^m V_{ij} + \dots + V_{1\dots m} \quad (\text{II.8})$$

On décrit $V(E(Y|x_i))$ comme la variance partielle de Y produite uniquement par les variations x_i et on obtient ainsi les relations suivantes :

$$\begin{aligned} V &= V(Y) \\ V_i &= V(E(Y|x_i)) \\ V_{ij}(x_i, x_j) &= V(E(Y|x_i, x_j)) - V_i - V_j \\ V_{ijk}(x_i, x_j, x_k) &= V(E(Y|x_i, x_j, x_k)) - V_i - V_j - V_k - V_{ij} - V_{ik} - V_{jk} \\ &\dots \\ V_{i\dots m}(x_i, \dots, x_m) &= V(E(Y|x_i, \dots, x_m)) - \sum_{i=1}^m V_i - \sum_{1 \leq i < j \leq m} V_{ij} - \dots - \sum_{i=1}^m V_{1\dots i-1, i+1\dots m} \end{aligned} \quad (\text{II.9})$$

Il est alors possible de déterminer l'effet d'une variable comme la part de variance produite par cette variable sur la variance totale. En ne considérant que l'effet seul de

la variable, on obtient l'indice de sensibilité du premier ordre qui mesure la part de variance de sortie induite par la seule variable x_i . Cet indice est également appelé *indice de Sobol du premier ordre* (Sobol', 1990) ou effet principal et correspond à un taux de corrélation.

$$S_i = \frac{V_i}{V} \quad (\text{II.10})$$

Les autres variances $V_{ij\dots}$ sont associées aux effets produits par des variations communes de plusieurs entrées. Ce sont des indices de sensibilité d'ordre supérieur allant de 2 à m notés S_{ij} à $S_{1\dots m}$ et ont la même définition que (II.10). On les nomme *indices de Sobol d'ordre supérieur* ou effets par interaction.

A partir de l'équation (II.8), on en déduit que la somme de tous les effets est égale à un et l'intégralité des variations induites et donc de la variance est étudiée :

$$1 = \sum_{i=1}^m S_i + \sum_{i<j}^m S_{ij} + \dots + S_{1\dots m} \quad (\text{II.11})$$

Pour déterminer toutes les contributions de la variable x_i , l'indice de sensibilité total est défini. Il est évalué par la somme de tous les indices faisant intervenir la i -ème variable seule et par interaction :

$$ST_i = S_i + \sum_{j \neq i}^m S_{ij} + \sum_{j < k | j, k \neq i}^m S_{ijk} + \dots + S_{1\dots i\dots m} \quad (\text{II.12})$$

A partir des équations (II.11) et (II.12), deux relations importantes sont déduites pour l'interprétation des indices qui sera faite par la suite.

$$\begin{aligned} \sum_{i=1}^m S_i &\leq 1 \\ \sum_{i=1}^m ST_i &\geq 1 \end{aligned} \quad (\text{II.13})$$

D'autres indicateurs peuvent être utilisés pour exprimer l'effet des variables. Le choix va dépendre de l'utilisation qui en est faite, certains indicateurs notamment ont un comportement logarithmique pour mettre en avant les variables influentes. Les indices de Sobol' ont été choisis, car ils offrent la représentation la plus claire. Les indices sont normés, si la valeur de l'indice associé à une variable est un, alors la variable explique toutes les variations de sortie, si la valeur est zéro la variable n'a pas d'effet et entre les deux l'interprétation de l'effet est linéaire. De plus, si la somme de tous les effets est différente de un cela dénote la présence d'interactions

entre les variables –voir équation (II.13). Cette représentation est largement utilisée en analyse de sensibilité et sera employée dans ce travail pour quantifier les effets dans le cadre de la simulation thermique dynamique.

Les indices de sensibilité sont donc obtenus à partir de différentes évaluations de la variance. Dans la suite, trois méthodes sont présentées pour obtenir ces résultats. Les deux premières sont définies dans ce chapitre, l'une est basée sur l'étude directe de la variance et l'autre sur les transformées de Fourier. La troisième méthode sera présentée dans le Chapitre IV, car issu d'un méta modèle utilisé à d'autres fins. Les versions initiales de ces méthodes ne sont pas intégrées dans le programme détaillé en Chapitre III, mais il a été choisi de les présenter pour leur originalité, car elles offrent des évolutions très intéressantes qui ont été intégrées et qui seront décrites en section 3.

2.2 Evaluation de la variance – Méthode de SOBOL

Le calcul de la variance totale déterminée à l'équation (II.6), requiert un estimateur pour obtenir de la forme intégrale une forme discrétisée. A partir d'un échantillon stochastique de N essais, suffisamment grand pour permettre une évaluation correcte, les estimateurs classiques pour calculer l'espérance et la variance sont :

$$\begin{aligned}\hat{E} &= \frac{1}{N} \sum_{n=1}^N \mathcal{M}(x_1^n, x_2^n, \dots, x_m^n) \\ \hat{V} &= \frac{1}{N-1} \sum_{n=1}^N \left(\mathcal{M}(x_1^n, x_2^n, \dots, x_m^n)^2 - \hat{E}^2 \right)\end{aligned}\tag{II.14}$$

Pour donner un ordre de grandeur, la taille de l'échantillon généralement utilisé pour que ces estimateurs fournissent une estimation correcte est comprise entre quelques centaines et un millier (Saltelli et al., 2004).

L'évaluation des variances partielles qui sont des variances conditionnelles telles que décrites en (II.9) demande une méthode pour isoler l'effet d'une ou plusieurs variables. La variance partielle V_i doit être déterminée à partir de la variance produite par la seule variable x_i sans ses interactions et quel que soient les valeurs des autres variables.

On produit deux échantillonnages **A** et **B** ayant des variations aléatoires, mais avec les mêmes propriétés statistiques. Ceux-ci sont représentés sous forme de matrice où chaque ligne correspond à une simulation :

$$\mathbf{A} = \begin{matrix} & \begin{matrix} x_1 & x_2 & \cdots & x_m \end{matrix} \\ \begin{matrix} a_1^1 \\ a_1^2 \\ \vdots \\ a_1^N \end{matrix} & \begin{matrix} a_2^1 \\ a_2^2 \\ \vdots \\ a_2^N \end{matrix} & \cdots & \begin{matrix} a_m^1 \\ a_m^2 \\ \vdots \\ a_m^N \end{matrix} \end{matrix}; \quad \mathbf{B} = \begin{matrix} & \begin{matrix} x_1 & x_2 & \cdots & x_m \end{matrix} \\ \begin{matrix} b_1^1 \\ b_1^2 \\ \vdots \\ b_1^N \end{matrix} & \begin{matrix} b_2^1 \\ b_2^2 \\ \vdots \\ b_2^N \end{matrix} & \cdots & \begin{matrix} b_m^1 \\ b_m^2 \\ \vdots \\ b_m^N \end{matrix} \end{matrix} \quad (\text{II.15})$$

Afin d'isoler l'effet de la variable x_i , on remplace dans \mathbf{B} les variations de la i -ème colonne par les variations de la i -ème colonne de \mathbf{A} . On construit la matrice \mathbf{B}_i . Cet échantillonnage \mathbf{B}_i est donc aléatoire par rapport à \mathbf{A} sauf pour la i -ème variable qui est la seule inchangée telle qu'on note la matrice :

$$\mathbf{B}_i = \begin{matrix} & \begin{matrix} x_1 & \cdots & x_i & \cdots & x_m \end{matrix} \\ \begin{matrix} b_1^1 \\ b_1^2 \\ \vdots \\ b_1^N \end{matrix} & \cdots & \begin{matrix} a_i^1 \\ a_i^2 \\ \vdots \\ a_i^N \end{matrix} & \cdots & \begin{matrix} b_m^1 \\ b_m^2 \\ \vdots \\ b_m^N \end{matrix} \end{matrix} \quad (\text{II.16})$$

La variance conditionnelle associée à la i -ème variable est alors donnée par (II.17), qui par le produit des résultats de ces deux échantillonnages identifie l'effet de la variable fixée. Ce résultat est ensuite directement utilisé dans l'estimateur de la variance précédemment décrit pour calculer la variance conditionnelle de la variable i .

$$\hat{V}_i = \frac{1}{N-1} \sum_{n=1}^N \left(\mathcal{M}(\mathbf{A}^{(n)}) \cdot \mathcal{M}(\mathbf{B}_i^{(n)}) - \hat{E}^2 \right) \quad (\text{II.17})$$

Le calcul de tous les indices du premier ordre demande alors de réaliser un jeu de simulation de référence \mathbf{A} , ainsi qu'un jeu de simulation \mathbf{B}_i pour chaque variable soit $(1+m)N$ simulations.

Le calcul des indices d'ordre supérieur suit la même démarche à la différence qu'il est nécessaire d'identifier l'effet conjoint de plusieurs variables. Pour l'ordre 2 on construit un échantillon \mathbf{B}_{ij} dans lequel les deux colonnes i et j ont été modifiées. L'étude de la variance avec cette condition contient ainsi la variance produite par x_i , x_j et leur interaction. Il est donc nécessaire de retrancher les variances d'ordre inférieur pour obtenir l'interaction désirée :

$$\begin{aligned}
V\left(E\left(Y|x_i, x_j\right)\right) &= V_{ij} + V_i + V_j \\
\hat{V}_{ij} &= \frac{1}{N-1} \sum_{n=1}^N \left(\mathcal{M}\left(\mathbf{A}^{(n)}\right) \cdot \mathcal{M}\left(\mathbf{B}_{ij}^{(n)}\right) - \hat{E}^2 \right) - \hat{V}_i - \hat{V}_j
\end{aligned}
\tag{II.18}$$

Ce calcul des interactions d'ordre 2 nécessite de réaliser $m \cdot (m-1) \cdot N/2$ simulations supplémentaires. La démarche est ensuite similaire pour les ordres supérieurs. Le nombre de combinaisons possible à chaque ordre est donné par les coefficients binomiaux. L'étude du modèle à l'ordre p peut être réalisé pour un nombre total de simulations tel que :

$$\begin{aligned}
N_{tot} &= \sum_{i=1}^p C_m^i \cdot N \leq 2^m \cdot N \\
&= \sum_{i=1}^p \frac{m!}{i! \cdot (m-i)!} \cdot N
\end{aligned}
\tag{II.19}$$

Le nombre de simulations nécessaire augmente alors de façon très importante dès qu'il est question d'étudier les interactions. Une alternative est d'estimer directement la somme de toutes les interactions de chaque variable avec le calcul des effets totaux puis d'analyser les résultats comme présenté à l'équation (II.13). La différence entre les effets principaux et totaux informe alors sur la présence d'interactions.

La variance produite par une variable avec toutes ces interactions n'est cependant pas directement accessible. Observer une interaction implique de fixer toutes les variables qui la compose, cependant l'estimation de la variance conditionnelle qui est faite avec cet essai contient les variances d'ordre inférieur de toutes les combinaisons des variables fixées.

$$V\left(E\left(Y|x_i, x_j, x_k\right)\right) = V_i + V_j + V_k + V_{ij} + V_{ik} + V_{jk} + V_{ijk} \tag{II.20}$$

Il est au contraire possible de calculer la variance de toutes les entrées sauf celle qui nous intéresse. On écrit ainsi :

$$\begin{aligned}
V_{-i} &= V\left(E\left(Y|\mathbf{x} \setminus x_i\right)\right) = V\left(E\left(Y|x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_m\right)\right) \\
&= \sum_{\substack{j=1 \\ j \neq i}}^m V_j + \sum_{\substack{1 \leq j < k \leq m \\ j \neq i, k \neq i}} V_{jk} + \dots + V_{1 \dots i-1, i+1 \dots m}
\end{aligned}
\tag{II.21}$$

La variance conditionnelle obtenue contient alors l'intégralité des effets non désirés. La différence avec la variance totale est alors uniquement due à la variable qui n'a

pas changé. On obtient ainsi la variance produite par la variable avec toutes ces interactions incluses (Saltelli et al., 2010).

$$V - V_{-i} = V_i + \sum_{\substack{1 \leq j \leq m \\ j \neq i}} V_{ij} + \dots + V_{1\dots m} \quad (\text{II.22})$$

Ce qui permet de former l'indice de sensibilité total associé à la variable :

$$ST_i = \frac{V - V_{-i}}{V} \quad (\text{II.23})$$

Pour réaliser ce calcul on produit un autre type d'échantillonnage \mathbf{B}_{-i} dans lequel on ne laisse que la i -ème colonne inchangée (Homma et Saltelli, 1996).

$$\mathbf{B}_{-i} = \begin{matrix} & x_1 & \dots & x_i & \dots & x_m \\ \begin{pmatrix} a_1^1 & \dots & b_i^1 & \dots & a_m^1 \\ a_1^2 & \dots & b_i^2 & \dots & a_m^2 \\ \vdots & & \vdots & & \vdots \\ a_1^N & \dots & b_i^N & \dots & a_m^N \end{pmatrix} \end{matrix} \quad (\text{II.24})$$

Puis la variance conditionnelle associée est calculée.

$$\hat{V}_{-i} = \frac{1}{N-1} \sum_{n=1}^N \left(\mathcal{M}(\mathbf{A}^{(n)}) \cdot \mathcal{M}(\mathbf{B}_{-i}^{(n)}) - \hat{E}^2 \right) \quad (\text{II.25})$$

Ce nouveau calcul ne permet pas de différencier les interactions, mais permet de connaître précisément leurs ordres de grandeur pour chaque variable avec $m \cdot N$ simulation supplémentaires.

La problématique du coût de simulation pour une étude plus détaillée de ces interactions a été étudiée par A. Saltelli. Une fois les évaluations pour l'étude de tous les effets principaux et de tous les effets totaux réalisées, celui-ci propose une méthode pour croiser ces évaluations afin d'évaluer l'effet total commun de deux facteurs et l'effet total commun $(m-2)$ facteurs (Saltelli, 2002). Cette technique prend généralement le nom de Sobol-Saltelli.

2.3 Techniques fréquentielles – FAST

La méthode FAST (Fourier Amplitude Sensitivity Test) est plus ancienne, elle a été introduite par Cukier et al. en 1973 (Cukier, 1973; Cukier et al., 1975; Schaibly, 1973) et a servi de base à la technique de Sobol précédemment décrite. Elle se base sur le

fait que si une variable d'entrée est influente, la réalisation de N simulations avec une variation sinusoïdale de la valeur d'entrée au cours des simulations produira N résultats de la sortie avec une variation sinusoïdale de même fréquence. L'amplitude de la variation de sortie informera alors sur l'influence plus ou moins importante de la variable au sein du modèle. En donnant une fréquence particulière à chaque variable il est possible d'évaluer la participation de chacune d'elle par une analyse fréquentielle complète de la sortie. Cette méthode permet donc l'estimation de façon simultanée de tous les indices à l'aide de la projection dans les bases de Fourier.

Cette approche repose sur la théorie ergodique qui admet qu'une évolution suffisamment longue d'un signal au cours du temps apporte les mêmes informations statistiques qu'un processus stochastique (Weyl, 1938). Le calcul de l'intégrale de dimension m nécessaire pour l'évaluation de la variance sur le domaine D_x peut donc être réalisé par une intégrale simple dépendante de l'évolution des essais dans le temps. Il est ainsi possible d'avoir une représentation similaire aux indices de Sobol' précédemment décrits. Cela demande de redéfinir l'échantillonnage des entrées en fonction de l'évolution d'une unique variable s qui permet l'exploration du domaine D_x . Une fréquence particulière est attribuée à chaque entrée afin de réaliser une variation simultanée selon la variable s et qui permet alors à chaque variable de balayer son espace de définition :

$$x_i^n = G_i(\sin(\omega_i \cdot s)) \quad (\text{II.26})$$

Afin d'intégrer l'analyse d'incertitude à cette analyse de sensibilité il est nécessaire d'assurer lors de l'exploration de l'espace le bon respect de la distribution d'origine de chaque variable. Au cours de l'évolution des simulations les variations d'entrée sont générées à partir de sinusoïdes. La fonction G_i qui apparaît dans l'équation (II.26) permet de modifier la distribution du signal sinusoïdal et la plage de variation. Le choix de la fonction $G_i(u)$ doit alors satisfaire la relation suivante :

$$\pi \cdot \sqrt{1-u^2} \cdot p_i \cdot \frac{dG_i(u)}{du} = 1 \quad (\text{II.27})$$

où p_i est la densité de probabilité de la variable considérée. On donne ainsi pour chaque variable un échantillonnage particulier en lui attribuant les bonnes propriétés de variation et de dispersion. Par exemple pour un échantillonnage uniforme entre $[-1;1]$ la fonction est de la forme $G(u) = 2 \cdot \arcsin(u)/\pi$ ce qui transforme de signal sinusoïdal en signal triangle. La figure ci-dessous présente les dispersions pour le signal u sinusoïdal et le résultat $G(u)$ triangulaire.

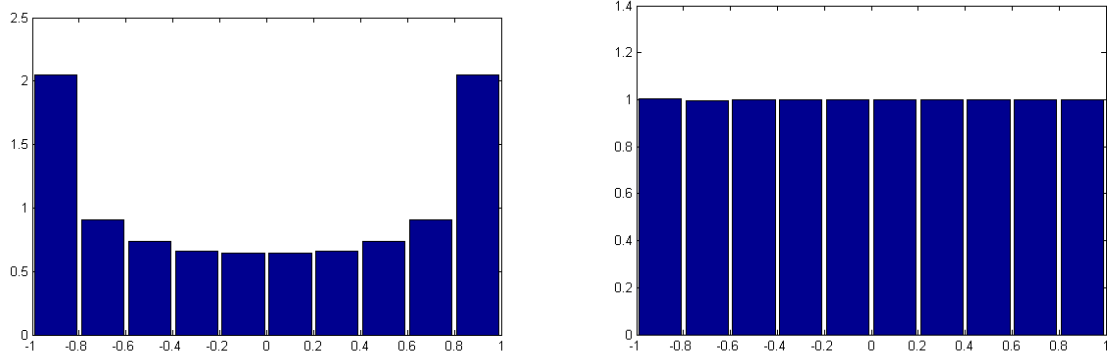


Figure II-1 : Exemple de dispersion pour un signal sinusoïdal puis résultat après traitement

D'autre part, le choix des fréquences est une phase importante qui détermine la capacité à observer et différencier les effets des différentes variables et de leurs interactions. Afin de permettre une étude correcte, tous ces effets doivent être associés à des fréquences différentes. Cela s'exprime par une condition d'incommensurabilité, c'est-à-dire qu'aucune fréquence ne doit être une combinaison linéaire des autres. Le jeu de fréquence choisi est incommensurable s'il ne vérifie la relation (II.28) qu'à la condition que $a_i = 0$ quel que soit i :

$$\sum_{i=1}^m a_i \omega_i = 0 \quad , a_i \in \mathbb{N} \quad (\text{II.28})$$

Le choix d'un jeu incommensurable de fréquence implique qu'il n'y a pas de périodicité dans l'exploration de l'espace. L'utilisation de fréquences non entières permet de respecter facilement la condition. La première illustration de la Figure II-2 montre l'évolution de deux variables pourvues des fréquences $\omega_1 = \sqrt{5}$ et $\omega_2 = 10\sqrt{3}$ incommensurables. La transformation ergodique permet alors d'écrire pour le calcul de la moyenne :

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \mathcal{M}(\mathbf{x}(s)) ds = \int_{D_x} \mathcal{M}(\mathbf{x}) p_x dx_1 \dots dx_m = f_0 \quad (\text{II.29})$$

Ce calcul requiert en théorie un nombre infini d'essais afin d'explorer l'intégralité de l'espace au bout d'un temps suffisamment long. La moyenne temporelle se confond ainsi avec la moyenne spatiale. Cependant, pour les besoins de l'analyse le choix des fréquences doit être limité aux nombres entiers et l'équation (II.28) est obligatoirement validée. La seconde illustration de la Figure II-2 décrit l'évolution des points avec $\omega_1 = 5$ et $\omega_2 = 9$ l'exploration du domaine est alors 2π périodique, mais ne couvre plus l'intégralité de l'espace. Les fréquences pouvant être une combinaison linéaire, le choix de celles-ci doit être fait avec attention et de manière à

ce que les combinaisons possibles tombent sur les harmoniques de haute fréquence afin de ne pas perturber l'analyse des effets.

La variable s explore donc l'espace de façon 2π -périodique et le domaine d'étude peut être réduit. L'intégrale fournit alors un estimateur de la moyenne :

$$\hat{f}_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{M}(\mathbf{x}(s)) ds \quad (\text{II.30})$$

Pour la variance l'estimateur est de la forme suivante :

$$\hat{V} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{M}(\mathbf{x}(s))^2 ds - \hat{f}_0^2 \quad (\text{II.31})$$

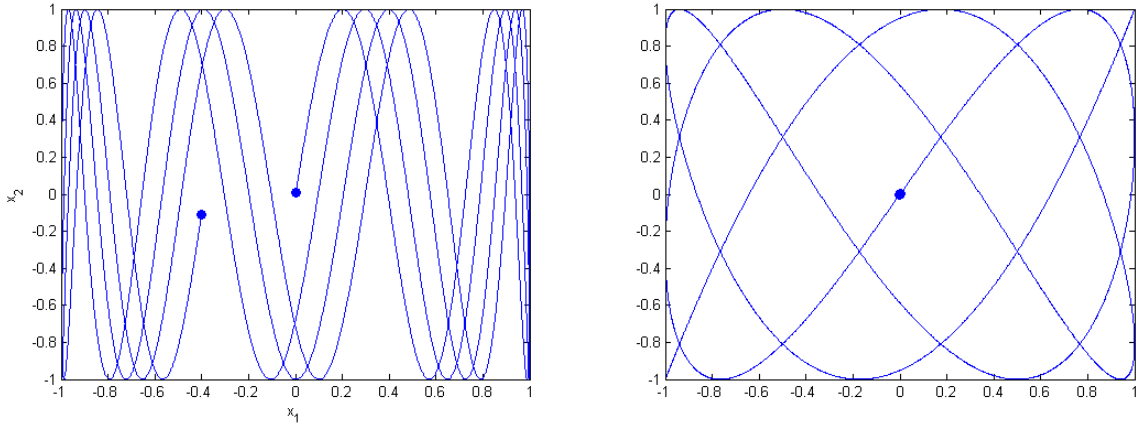


Figure II-2 : Evolution d'un couple de variables générées à partir de fréquences incommensurables puis de fréquences entières

A partir de là, les équations générales du traitement des signaux peuvent être employées. L'application de l'égalité de Parseval permet de passer de la forme intégrale à une somme de coefficient de Fourier. La variance totale est la somme des tous les coefficients sans la moyenne :

$$\hat{V} = \sum_{\substack{j=-\infty \\ j \neq 0}}^{+\infty} |c_j(\mathcal{M})|^2 \quad (\text{II.32})$$

c_j est défini à partir des coefficients classiques du développement en série de Fourier du signal en cosinus et sinus tel que :

$$c_j = \begin{cases} a_j - i \cdot b_j & ; j \geq 0 \\ a_j + i \cdot b_j & ; j < 0 \end{cases}$$

$$a_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{M}(\mathbf{x}(s)) \cos(js) ds$$

$$b_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{M}(\mathbf{x}(s)) \sin(js) ds$$
(II.33)

En pratique la décomposition est obtenue à partir d'un nombre fini d'essais et donne ainsi des estimateurs des coefficients de Fourier.

$$s_n = 2\pi \cdot \frac{n}{N}$$
(II.34)

$$\hat{a}_j = \frac{1}{N-1} \sum_{n=1}^N \mathcal{M}(\mathbf{x}(s_n)) \cdot \cos(js_n) \quad a_j = a_{-j}$$

$$\hat{b}_j = \frac{1}{N-1} \sum_{n=1}^N \mathcal{M}(\mathbf{x}(s_n)) \cdot \sin(js_n) \quad b_j = b_{-j}$$
(II.35)

Ces estimateurs ont été normalisés par $N-1$ afin de correspondre à la définition de la variance non biaisée qui est exprimée par l'équation (II.14). La forme finale de l'estimateur de la variance définie à l'équation (II.32) est donc :

$$\hat{V} = 2 \sum_{j=1}^{(N-1)/2} |c_j|^2$$
(II.36)

On remarque que l'indice j ne dépasse pas la valeur de $(N-1)/2$ à cause de la symétrie des coefficients de Fourier, il sera donc impossible d'évaluer les fréquences supérieures à cette valeur. Cette limite correspond à la condition fondamentale du théorème d'échantillonnage qui détermine, en fonction du nombre d'essais, la plus haute fréquence observable. Mais pour une étude correcte il est nécessaire d'inclure dans le calcul un certain nombre d'harmoniques supplémentaires même si celles-ci contiennent généralement peu de puissance. On définit alors un facteur d'interférence M , prenant une valeur de 4 ou 6 en fonction des auteurs, qui fixe le nombre d'essai nécessaire selon la fréquence maximale d'échantillonnage :

$$N = 2 \cdot M \cdot \omega_{\max} + 1$$
(II.37)

Pour obtenir une variance partielle et déterminer l'effet dû à une seule variable, une fraction de la variance totale doit être identifiée. Etant donné que chaque entrée est associée à une pulsation, la variance partielle de la variable x_i est alors la somme de tous les coefficients de Fourier faisant intervenir sa pulsation ω_i et les harmoniques

associées. Par exemple, si dans le modèle x_i est élevé au carré on produit une interaction (x_i, x_i) et deux contributions apparaissent sur $\{0; 2\omega_i\}$. Les indices de sensibilité sont alors déterminés par :

$$S_i = \frac{\sum_{j=1/\omega=k\omega_i}^{N/2} |c_j|^2}{\sum_{j=1}^{N/2} |c_j|^2} \quad (\text{II.38})$$

En ce qui concerne les interactions, il est possible d'étendre l'analyse pour réaliser cette étude. Si deux variables sont échantillonnées à des pulsations ω_1 et ω_2 , l'interaction entre celles-ci apparaît en sortie sur les fréquences $(\omega_1 \pm \omega_2)$, mais également sur toutes les combinaisons observables de ces signaux avec leurs harmoniques.

Le choix des fréquences est donc particulièrement important pour le bon déroulement de l'analyse. En fonction des besoins de l'étude il peut être nécessaire d'augmenter de façon importante les fréquences, et donc le nombre de simulations pour que ces différentes combinaisons ne se confondent pas. Cette étape est donc la plus délicate, car elle impacte directement sur la performance de la méthode et la qualité des résultats. Certains auteurs ont développé des tables optimales pour permettre un choix plus aisé de ses fréquences (Schaibly, 1973).

Au final, FAST est une méthode plus performante que celle de Sobol pour l'évaluation des effets principaux, le calcul se fait avec une précision équivalente mais avec un nombre de simulations plus faible (Mara et Rakoto Joseph, 2008). Par contre celle-ci est mise en défaut pour les ordres supérieurs, car ces effets sont généralement faibles et dispersés sur de nombreuses fréquences ce qui les rend particulièrement difficiles à évaluer.

La méthode a par la suite connu des modifications notamment avec EFAST –*Extended FAST*– (Saltelli et al., 1999) qui attribue des fréquences variables aux entrées et RBD-FAST (Tarantola et al., 2006) qui est présenté dans la suite du chapitre. Cette dernière méthode présente un intérêt pour la réduction du nombre de simulations et cela également pour la méthode d'I.M. Sobol.

3 Améliorations pour la réduction du coût de calcul – RBD

Les méthodes précédentes Sobol et FAST, permettent de déterminer avec précision les principales relations au sein du modèle. Dans la plupart des modèles la quasi-intégralité des variations sont expliquées par les effets principaux. L'évaluation des

autres effets est alors réduite à une étude générale de la variance résiduelle, c'est-à-dire la part de variance qui n'a pas pu être attribuée aux effets principaux. Si cette valeur est suffisamment faible les effets par interaction sont jugés non-influents et se perdront dans l'incertitude associée aux résultats. L'évaluation de ces effets est cependant réalisable, mais au prix d'un nombre d'essais très importants afin d'obtenir assez de précision pour différencier ces effets secondaires. L'utilité de ces nouvelles informations ne justifie alors généralement pas le coût nécessaire pour cette étude. Les effets secondaires sont par conséquent très rarement calculés.

Si on fait l'hypothèse que seuls les effets principaux sont nécessaires, il est possible d'améliorer l'efficacité de ces méthodes. Réduire l'étude à la seule obtention des effets principaux permet de modifier les techniques afin d'améliorer leur capacité à extraire les informations demandées avec un jeu réduit de simulation. Cette réduction d'information permet tout de même de réaliser une évaluation globale de la présence d'interactions, nous reviendrons en Chapitre IV :3.3 sur l'étude particulière de ces interactions via un méta modèle.

L'objectif de cette partie est de permettre d'extraire les informations désirées avec la précision voulue et le moins d'appel possible au modèle. Nous présentons pour cela la technique RBD (*Random Balance Design*) décrite par Satterthwaite en 1959. Il s'agit d'une technique employée dans le milieu industriel depuis 1956 pour l'application de méthodes statistiques basées sur les plans d'expériences (Taguchi, 1986). Dans les problèmes complexes, réaliser une étude classique via les plans d'expériences demande de réaliser un nombre d'essais très important. En voulant réduire le nombre d'essais on passe d'un plan complet, à un plan saturé, puis à un plan sursaturé. C'est-à-dire que le nombre de facteur dépasse le nombre d'essai, ce qui implique des hypothèses sur le fait de négliger ou non certains effets. D'autre part l'équilibre parfait du plan n'est plus forcément assuré, car le plan n'est plus symétrique vis-à-vis de chaque variable.

La mise en place de ce type de plan devient alors subjective, car elle repose sur des hypothèses *a priori* qui ne sont pas directement vérifiables et qui nécessitent des expériences supplémentaires. Dans ce cas certaines études sont réalisées avec des plans à équilibre aléatoire tels que décrits par Satterthwaite (Satterthwaite, 1959). Le principe est de mélanger l'ordre des variables d'entrée avant les simulations afin de rendre aléatoire l'exploration du domaine tout en gardant les bonnes propriétés statistiques au niveau de chaque variable et permettre de réaliser une estimation correcte de chaque effet. Le passage à un plan aléatoire permet alors de diluer les erreurs dans la variance (Booth et Cox, 1962; Morris, 1991) et permet de trouver

rapidement les principaux facteurs influents avec un coût de simulation acceptable et sans faire d'hypothèses sur le modèle.

La méthode a largement été critiquée, mais elle trouve des applications très intéressantes dans l'analyse de sensibilité. Le principe a été repris et étendu à l'analyse de sensibilité globale (Tarantola et al., 2006). Dans les sections suivantes nous présentons l'application de ce principe aux deux méthodes précédemment décrites ainsi que quelques modifications qui améliorent les résultats. La dernière section présente une méthode d'évaluation de l'incertitude présente dans l'évaluation des indices de sensibilité.

3.1 SOBOL' et améliorations

3.1.1 Méthode RBD-Sobol

Pour détailler l'application de la méthode RBD à l'analyse globale il est nécessaire revenir à la définition des indices de Sobol faite en section 2.2 ci-dessus. Le calcul des indices de sensibilité via cette méthode nécessite la réalisation d'un jeu de simulation de référence \mathbf{A} , puis m jeux de simulation modifiés pour l'étude des m variables d'entrée. Chacun de ces jeux de simulation modifiés \mathbf{B}_i sont construits de telle sorte que seule la variable x_i est identique au jeu de référence alors que toutes les autres sont aléatoires. Le nombre de simulations nécessaire est donc de $N_{total} = (m+1)N$. La méthode RBD permet de regrouper tous les seconds jeux de simulation \mathbf{B}_i dans un unique jeu \mathbf{B} . L'étude des effets principaux passe alors à seulement $N_{total} = 2N$ simulations et n'est *a priori* plus dépendant du nombre de variables étudiés.

Dans l'échantillonnage qui doit être réalisé, toutes les matrices \mathbf{B}_i doivent être semblables et construites à partir du même échantillonnage qui a permis de réaliser la matrice \mathbf{A} présentée à l'équation (II.15). A partir de la suite $[1..N]$, on produit pour chaque variable i une permutation circulaire σ_i indépendante :

$$\boldsymbol{\sigma} = \{\sigma_1, \sigma_2, \dots, \sigma_m\} \quad (\text{II.39})$$

Ces vecteurs forment une matrice de permutation et pour plus de détails voir la section 4.3 sur les méthodes d'échantillonnages. En appliquant ces permutations à chaque colonne de l'échantillonnage initial \mathbf{A} , une version aléatoire que l'on nomme \mathbf{B} est produite tout en conservant les mêmes propriétés statistiques que \mathbf{A} . Cette modification est réalisée par les vecteurs $\boldsymbol{\sigma}_i$ telle que $\mathbf{a}_1(\boldsymbol{\sigma}_1)$ est la réorganisation de \mathbf{a}_1 selon $\boldsymbol{\sigma}_1$:

$$\mathbf{A} = \begin{pmatrix} x_1 & x_2 & \cdots & x_m \\ a_1^1 & a_2^1 & \cdots & a_m^1 \\ a_1^2 & a_2^2 & \cdots & a_m^2 \\ \vdots & \vdots & \vdots & \vdots \\ a_1^N & a_2^N & \cdots & a_m^N \end{pmatrix} \quad (II.40)$$

$$\mathbf{B} = \begin{pmatrix} b_1^1 & b_2^1 & \cdots & b_m^1 \\ b_1^2 & b_2^2 & \cdots & b_m^2 \\ \vdots & \vdots & \vdots & \vdots \\ b_1^N & b_2^N & \cdots & b_m^N \end{pmatrix} = [\mathbf{a}_1(\sigma_1) \quad \mathbf{a}_2(\sigma_2) \quad \cdots \quad \mathbf{a}_m(\sigma_m)]$$

Ces deux échantillonnages correspondent aux $2N$ simulations nécessaires pour estimer les effets principaux. Chaque ligne correspond à un essai et un résultat y est associé. Afin de conserver l'intégrité des résultats les prochaines permutations devront donc être appliquées sur toutes les colonnes de chaque jeu ainsi que les résultats associés.

A partir des permutations de base (II.39), il est ensuite possible de produire les permutations inverses σ_{-i} de chaque variable afin de pouvoir retourner à l'état initial.

$$\mathbf{b}_i(\sigma_{-i}) = \mathbf{a}_i \quad (II.41)$$

Appliquer une de ces permutations à l'ensemble de \mathbf{B} permet de retrouver un arrangement similaire à \mathbf{B}_i c'est-à-dire une matrice qui a les mêmes variations que \mathbf{A} dans sa i -ème colonne. La colonne concernée par la permutation inverse retourne à son état initial, mais toutes les autres colonnes subissent une nouvelle permutation aléatoire.

$$\mathbf{B}(\sigma_{-i}) = [\mathbf{b}_1(\sigma_{-i}) \quad \cdots \quad \mathbf{a}_i \quad \cdots \quad \mathbf{b}_m(\sigma_{-i})] \sim \mathbf{B}_i \quad (II.42)$$

Le calcul des effets est alors inchangé et peut être réalisé comme précédemment, la notation vue à l'équation (II.17) est modifiée pour intégrer la permutation inverse σ_{-i} associée à chaque variable :

$$\hat{V}_i = \frac{1}{N-1} \sum_{n=1}^N \left(\mathcal{M}(\mathbf{A}^{(n)}) \cdot \mathcal{M}(\mathbf{B}(\sigma_{-i})^{(n)}) - \hat{E}^2 \right) \quad (II.43)$$

3.1.2 Modification des estimateurs de variance

La communauté de l'analyse de sensibilité s'efforce de développer des estimateurs de plus en plus robustes. L'utilisation de ces estimateurs permet une meilleure évaluation des indices de sensibilité. La liste suivante propose une amélioration de chacun des estimateurs précédemment décrits, mais bien d'autres existent (Saltelli et al., 2010).

L'espérance au carré peut être évaluée avec l'ensemble des essais (Homma et Saltelli, 1996) :

$$\hat{E}^2 = \frac{1}{N} \sum_{n=1}^N \mathcal{M}(\mathbf{A}^{(n)}) \cdot \mathcal{M}(\mathbf{B}^{(n)}) \quad (\text{II.44})$$

La variance peut également être évaluée avec l'ensemble des essais en posant :

$$\hat{V} = V \left(\frac{\mathcal{M}(\mathbf{A}) - \mathcal{M}(\mathbf{B})}{\sqrt{2}} \right) \quad (\text{II.45})$$

Pour le calcul de la variance conditionnelle, l'équation (II.44) conduit à poser (Sobol' et al., 2007) :

$$\hat{V}_i = \frac{1}{N} \sum_{n=1}^N \mathcal{M}(\mathbf{A}^{(n)}) \cdot (\mathcal{M}(\mathbf{B}_i^{(n)}) - \mathcal{M}(\mathbf{B}^{(n)})) \quad (\text{II.46})$$

Ainsi que :

$$\hat{V}_{-i} = \frac{1}{N} \sum_{n=1}^N \mathcal{M}(\mathbf{A}^{(n)}) \cdot (\mathcal{M}(\mathbf{A}^{(n)}) - \mathcal{M}(\mathbf{B}_{-i}^{(n)})) \quad (\text{II.47})$$

Tous ces estimateurs augmentent la précision des résultats, mais ne suffisent pas à assurer des résultats corrects à eux seuls. En effet la bonne évaluation des indices de sensibilité nécessite certes un bon estimateur, mais également un nombre d'essai suffisant, représentatif de la variabilité des entrées et avec une exploration multidimensionnelle de l'espace des variables d'entrée. La qualité des indices de sensibilité peut être évaluée via une technique présentée en section 3.3.

3.2 FAST et améliorations

3.2.1 Méthode RBD-FAST

Les permutations RBD peuvent également être appliquées à la méthode FAST pour former une extension appelée RBD-FAST. Elle a été développée par Tarantola et al.

(Tarantola et al., 2006) puis améliorée pour évaluer les indices de sensibilité globaux des entrées de façon indépendante ou par groupe (Mara, 2009).

La méthode FAST, qui a été présentée précédemment en section 2.3 ci-dessus, réalise l'étude des effets grâce à l'analyse fréquentielle des variations de sortie. Cela étant rendu possible par un comportement périodique des variables d'entrée. La principale difficulté réside dans le choix des fréquences attribuées à chaque entrée et d'autant plus si les interactions sont prises en compte. Cette technique dispose de variantes comme EFAST qui contourne cette difficulté avec une décomposition de l'étude variable par variable en attribuant successivement une fréquence basse à chaque variable alors que les autres variables ont des fréquences plus élevées. La démarche se rapproche de la méthode de Sobol dans le sens où chaque entrée est étudiée individuellement.

Avec RBD ce principe est étendu pour regrouper toutes ces simulations en un unique jeu. On réalise un échantillonnage de type FAST classique, mais avec une unique fréquence pour toutes les entrées. Le choix de la fréquence d'entrée n'est alors plus un paramètre influent et la pulsation peut être fixée à une valeur de un, ce qui correspond à une période sur les N simulations. Ensuite on réalise des permutations circulaires indépendantes sur chaque vecteur d'entrée comme montré pour la méthode de RBD-Sobol présentée en partie précédente à l'équation (II.39).

Comme pour la méthode FAST initiale, l'échantillonnage est créé à partir d'une variable d'exploration s qui permet de passer d'un problème à m dimensions à un problème unidimensionnel. Cette variable décrit chaque simulation de façon unique et va nous permettre d'organiser les simulations dans un ordre particulier en fonction des permutations aléatoires attribuées à chaque variable. A partir des équations (II.26) et (II.34) on déduit :

$$x_i^n = G_i \left(\sin \left(\omega_0 \cdot 2\pi \cdot \frac{\sigma_i(n)}{N} \right) \right), n \in [1, N] \quad (\text{II.48})$$

Chaque variable x_i décrit initialement une sinusoïde au cours des simulations. L'intégration de la permutation circulaire rend aléatoire ce comportement périodique tout en conservant les valeurs prises et les propriétés statistiques. L'échantillonnage ainsi obtenu présente une évolution aléatoire et indépendante pour chaque variable.

Dans un premier temps il est important de noter que quel que soit la forme du signal de sortie l'analyse fréquentielle est réalisable. Un signal aléatoire est un bruit blanc et est composé d'une infinité de fréquences. La décomposition en série de Fourier

répartit alors l'énergie sur le spectre étudié et la somme des coefficients calculée permet d'obtenir la même définition que la variance classique.

En ordonnant les essais avec la permutation inverse correspondante à la i -ème variable σ_{-i} , on retrouve la sinusoïde initiale pour cette variable, mais toutes les autres sont mélangées. On fait donc apparaître sur le spectre de la sortie un pic à la fréquence ω_0 et ses harmoniques dues à la variable étudiée. Toutes les autres variables sont organisées aléatoirement et produisent un bruit blanc.

On sait donc que l'effet de la i -ème variable se manifeste dans les basses fréquences et ses interactions sont une combinaison linéaire d'un bruit blanc, donc perdues avec les effets de toutes les autres variables. La puissance du signal diminue rapidement avec l'augmentation des harmoniques. D'après les différents contributeurs de FAST ou RBD-FAST 4 à 6 harmoniques sont généralement suffisantes pour identifier les effets d'une variable. La valeur M est alors définie comme le nombre d'harmonique prises en compte dans la réponse. L'indice de sensibilité de la sortie est évalué comme précédemment, mais avec les coefficients de Fourier $c_{\sigma_{-i}}$ calculés à partir du signal de sortie réorganisé. Si de plus ω_0 est défini égal à un, il est possible d'exprimer la variance conditionnelle \hat{V}_i comme la somme de tous les coefficients de Fourier de un jusqu'à $N_{\text{lim}} = 2M$ tel que le conçoit le théorème d'échantillonnage.

$$\begin{aligned} \hat{V}_i &= \sum_{j=1}^{N_{\text{lim}}} |c_j^{(\sigma_{-i})}|^2 \\ \hat{V} &= \sum_{j=1}^N |c_j^{(\sigma_{-i})}|^2 \end{aligned} \quad \hat{S}_i = \frac{\hat{V}_i}{\hat{V}} \quad (\text{II.49})$$

La variance totale \hat{V} , est la somme de tous les coefficients –sans la moyenne– c'est-à-dire la somme précédente plus les coefficients restant qui représentent la suite du signal blanc produit par les autre variables. Pour obtenir une représentation précise de ces signaux de faible puissance il est nécessaire de disposer d'un nombre suffisant de valeurs. On détermine alors un nombre minimum d'essais à réaliser :

$$N = 2 \cdot (M + L) \omega_0 \quad \omega_0 = 1 \quad (\text{II.50})$$

L , est une valeur arbitraire positive généralement $L \sim 100$. Elle peut être vue comme un facteur qui détermine la précision et permet de produire au moins cent coefficients supplémentaires pour le calcul de la variance du bruit blanc en plus des 4 à 6 coefficients précédents.

Au final en considérant des valeurs classiques, la méthode permet de trouver les variables influentes avec un minimum de 212 essais quel que soit le nombre d'entrées.

3.2.2 Réduction du biais

La variance des entrées non observées est dispersée sur toutes les fréquences sans exception. Ainsi lors du calcul de la variance conditionnelle présente dans les fréquences basses, une partie du bruit blanc est attribué faussement à la variance produite par l'entrée étudiée. Les indices de sensibilité sont alors tous légèrement surestimés. La Figure II-3 décrit cet effet sur une représentation schématique du spectre obtenu lors de l'analyse. Pour rappel le calcul de l'effet de sensibilité se fait par le ratio de l'énergie à gauche de la séparation sur l'énergie totale.

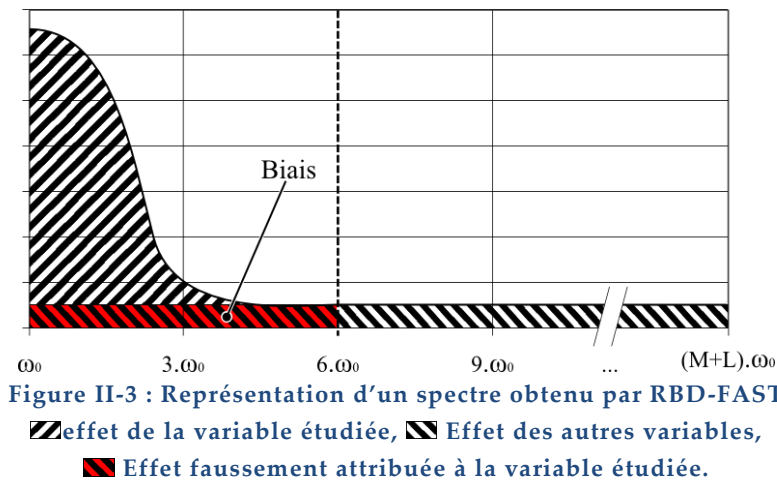


Figure II-3 : Représentation d'un spectre obtenu par RBD-FAST
 ▨ effet de la variable étudiée, ▩ Effet des autres variables,
 ■ Effet faussement attribué à la variable étudiée.

En considérant l'énergie du bruit blanc uniformément distribuée, ce biais est facilement pris en compte par l'ajout d'un terme correctif (Tissot et Prieur, 2012) :

$$\lambda = 2.N_{lim}/N$$

$$\hat{S}_i^c = \hat{S}_i - \frac{\lambda}{1-\lambda} (1 - \hat{S}_i) \quad (II.51)$$

La correction produite peut ainsi conduire à des indices de sensibilité négatifs dans les cas où l'effet de la variable est particulièrement faible. Cependant la méthode RBD-FAST est une analyse fréquentielle et les effets ne peuvent avoir que des valeurs positives. Cette valeur négative est due à l'erreur faite lors de l'estimation de la variance, elle donne ainsi une information grossière sur la précision des effets évalués.

Ce terme corrige ainsi la surestimation généralement constatée lors des comparatifs entre RBD-FAST et les autres techniques. Le résultat obtenu est alors de très bonne

qualité vis-à-vis du nombre de simulations réalisées, la Figure II-4 montre une convergence très rapide avec des résultats stables bien avant les 200 simulations. De plus la méthode n'est plus dépendante du nombre d'harmoniques M prises en compte et il est choisi d'augmenter cette valeur à 10 afin d'être sûr de ne pas couper une partie du signal.

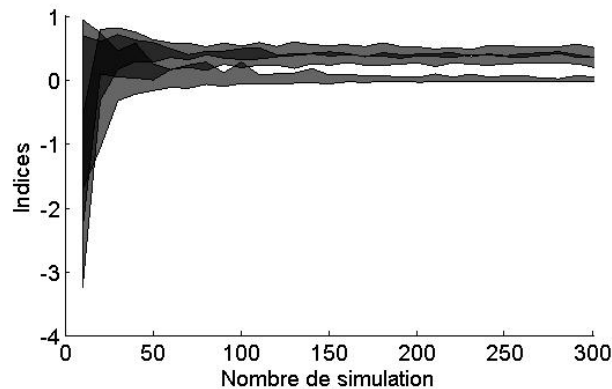


Figure II-4 : Convergence des indices de sensibilité sur la fonction d'Ishigami

3.2.3 Calcul sans discrétisation spécifique des entrées

Une évolution intéressante est la généralisation de la méthode RBD-FAST à tous type d'échantillonnages d'entrée. La permutation aléatoire des entrées permet de s'affranchir de l'importance de l'ordre des simulations et l'analyse est réalisable dès que l'on trouve une fréquence de variation basse dans une entrée alors que les autres se comportent aléatoirement.

L'évolution présentée consiste à reproduire une variation périodique sur une entrée via une organisation particulière des essais. A partir d'un échantillonnage aléatoire, les essais sont classés en fonction des valeurs croissantes d'une entrée. Les valeurs de cette entrée sont ensuite réorganisées alternativement en partant du centre tel que les pairs sont en première partie et les impairs sont en seconde ou inversement. La variable a alors une forme triangulaire de pulsation unitaire. La Figure II-5 présente l'évolution d'une variable d'entrée au cours de la mise en forme. On remarque ainsi qu'un échantillonnage uniforme formera un signal triangulaire tout comme FAST. Le dernier cadran de cette même figure représente le spectre obtenu après transformation et permet de se rendre compte de la faible puissance contenue dans les hautes harmoniques. Cette transformation n'impacte donc en aucun cas l'analyse qui sera faite par la suite.

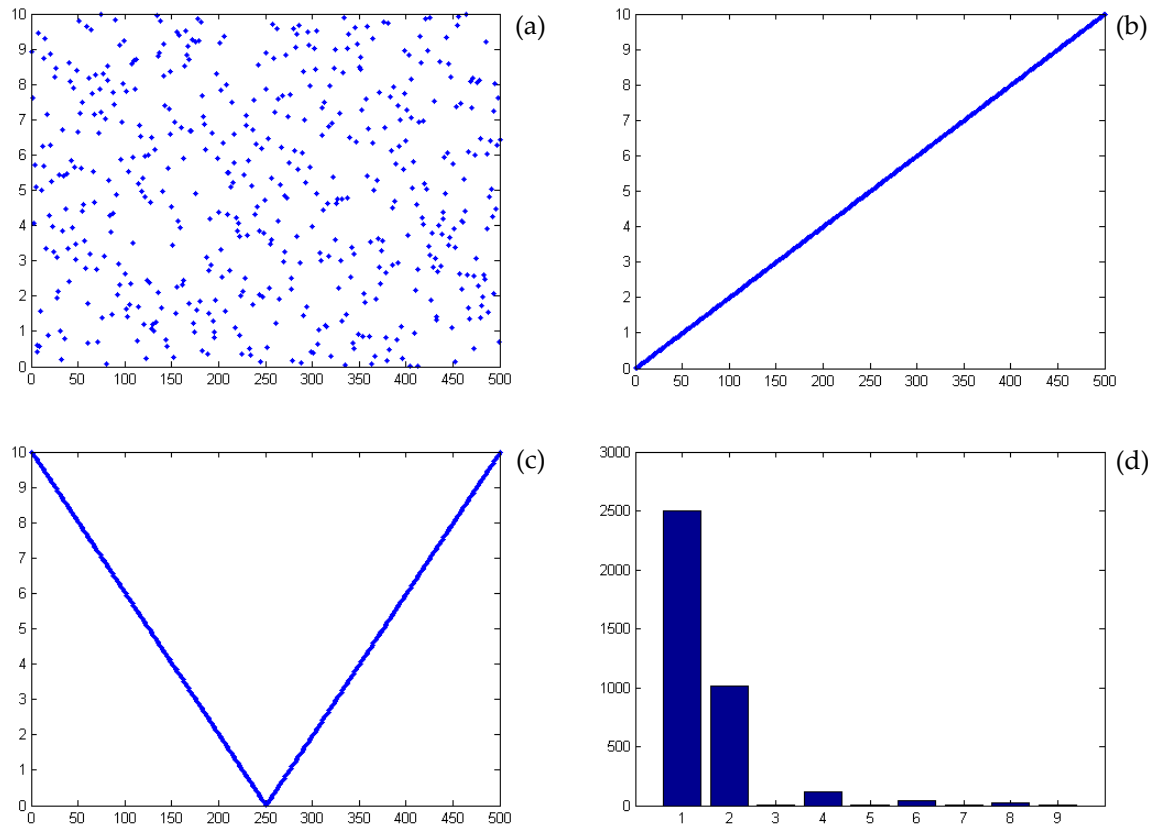


Figure II-5 : Construction d'un signal périodique à partir de variations aléatoires
 (a) échantillonnage aléatoire uniforme d'une variables d'entrée, (b) classement des essais selon les valeurs croissantes de cette entrée, (c) répartition alternative, (d) Spectre final

3.3 Estimateur de confiance – Bootstrap

Dans une démarche d'étude de l'incertitude il est important de fournir des évaluations fiables des effets des variables. L'incertitude de l'indicateur lui-même doit donc pouvoir être évaluée. La qualité d'un indice de sensibilité dépend de la méthode qui a servi à le produire, mais également du modèle qui a des effets et des interactions plus ou moins importants. De plus le nombre d'essai doit être suffisant pour l'évaluation et adapté au nombre de variable pour la bonne exploration de l'espace. Malgré l'utilisation d'estimateurs optimisés et dé-biaisés une réduction trop forte du nombre de simulations peut mener à des résultats incertains. Il est donc intéressant d'évaluer la qualité des indices produits.

Généralement chaque technique d'analyse intègre dans sa théorie une méthode d'évaluation spécifique de l'erreur. La mise en place de ces évaluations est alors contraignante et nécessite des simulations supplémentaires afin de réaliser une validation croisée pour ne pas évaluer l'erreur avec les mêmes points qui ont servi au calcul initial.

L'évaluation de la qualité de l'indice de sensibilité est alors souvent réalisée par une technique générique dite de Bootstrap. Elle évalue en fonction des données disponibles une dispersion probable des indices calculés. On évalue alors l'erreur, les intervalles de confiance, le biais et d'autre mesure de la précision (Efron, 1992).

Prenons pour exemple une population de $N = 500$ essais, qui à travers un processus étudié quelconque donne 62 résultats à 1 et les autres à 0. À partir de ces données une estimation de l'efficacité apparente du processus est réalisée et donne 12,4% de résultats favorables. Pour vérifier la validité de ce résultat nous procédons à un Bootstrap. Le calcul de l'estimation est refait selon une nouvelle population de même taille $N_{bootstrap} = 500$ tirée aléatoirement et avec remise à partir des essais initiaux. Ces résultats sont facilement obtenus car ils réutilisent les résultats des précédentes simulations et ne demandent donc pas de simulations supplémentaires. Cette population de Bootstrap nouvellement construite donne alors une information de l'efficacité du processus étudié certainement différente du résultat initial. En répétant cette évaluation aléatoire issues du Bootstrap de nombreuses fois, par exemple $N_{rep} = 1000$, on obtient assez de valeur pour avoir une représentation statistique cohérente.

Les différentes valeurs de Bootstrap produisent une dispersion supposée de forme gaussienne. En fonction des résultats et du nombre de simulations disponibles cette dispersion sera plus ou moins importante. L'évaluation de l'erreur est très proche de l'erreur commise par le processus stochastique initial et permet de mettre en avant un problème de stabilité de la solution. L'estimateur d'erreur peut alors être produit de différentes manières. Le calcul de l'écart type est la plus évidente et demande peu de répétitions. Un intervalle de confiance peut également être estimé sur le résultat en étudiant par exemple les 5^e et 95^e centiles. L'évaluation de ces centiles demande par contre un plus grand nombre de répétitions avec un minimum entre 500 et 1000 (Dubreuil et al., 2014). La position de ces limites permet alors de vérifier la symétrie de la distribution (Archer et al., 1997).

Ce calcul demande peu de ressources comparé au coût que représente l'appel au modèle énergétique, mais ne peut être considéré comme une réelle mesure de l'erreur. La méthode dispose cependant d'amélioration pour assurer une meilleure estimation, avec des techniques apparentées comme Bootstrap-after-Bootstrap ou Jackknife-After-Bootstrap mais dans notre cas ces informations ne sont pas nécessaires. L'écart entre les 5^e et 95^e centiles offre une estimation suffisante et permet une interprétation directe et générique de ce que peut être l'erreur commise sur les indicateurs.

4 Échantillonnage des entrées

4.1 L'importance du choix des évaluations

L'échantillonnage est le choix des points de l'espace des entrées du modèle qui seront évalués. Les techniques qui ont été précédemment décrites estiment l'effet des entrées par une analyse de la variance de sortie. Ce sont des méthodes qui considèrent le modèle comme une boîte noire et se basent donc uniquement sur les évaluations du modèle disponibles. Ainsi il est possible d'influencer l'observation du modèle qui est faite en fonction des points choisis.

L'échantillonnage doit permettre de réaliser une exploration cohérente de l'espace vis-à-vis du comportement que l'on souhaite étudier. Des méthodes d'échantillonnage ont été développées afin d'améliorer la représentativité et la qualité des points choisis pour l'évaluation du modèle.

Il est nécessaire de respecter deux propriétés importantes. D'une part, les entrées disposent de propriétés statistiques qui correspondent à l'objectif de l'évaluation et il est nécessaire de transmettre correctement ces propriétés via les variations des entrées. Les informations statistiques présentes dans ces variations ne doivent alors pas comporter de biais ce qui pourrait influencer les résultats avant même le début des simulations. Cette propriété est particulièrement mise en avant lorsque l'étude intègre des lois de probabilité avec des événements rares. D'autre part, il est important d'assurer via ces variations une exploration multidimensionnelle de l'espace qui permet de réaliser une exploitation uniforme – vis-à-vis des lois de distribution – et non corrélée de l'espace d'étude. Cette dernière propriété permet de plus d'assurer une convergence rapide des estimateurs utilisés.

4.2 Définition statistique des entrées

Les informations statistiques données aux entrées doivent être en accord avec l'objectif de l'étude. Le même processus d'évaluation peut avoir une finalité différente si les variations correspondent aux incertitudes d'entrées ou au contraire si elles correspondent à un comportement équiprobable pour évaluer uniquement les effets du modèle. On décrit ainsi les deux types d'analyse qui vont être faites.

L'une destinée à la recherche de solution et l'autre à la validation de l'état initial ou de l'état final. Dans tous les cas les entrées sont déterminées par des propriétés statistiques. Ce sont ces propriétés qui définiront les valeurs prises par les variables au cours des essais. Dans le cas actuel une définition sommaire des entrées sera faite.

On considère chaque entrée indépendante et définie via des lois de probabilité classiques.

Dans une recherche de solution la variation des entrées doit être homogène. Il n'est pas souhaitable de favoriser l'exploration du domaine sur une zone plutôt qu'une autre. Toutes les variations seront uniformes dans le domaine d'étude défini par les valeurs minimales et maximales des entrées choisies ou pour les données discrètes par une liste de choix possible dont la probabilité d'occurrence reste identique.

Une fois une solution choisie il est possible d'attribuer une nouvelle définition statistique aux variables d'entrée afin de prendre en compte l'incertitude présente sur les valeurs choisies. Les entrées discrètes sont alors fixes et les autres entrées peuvent être définies par un comportement uniforme ou gaussien. Ce nouveau comportement est défini par une moyenne et une plage de variation qui peut être fixe ($\pm 1\text{cm}$) ou relative par rapport à la moyenne ($\pm 10\%$). La définition d'une plage de variation est préférée par rapport à l'écart-type car la signification est plus facilement interprétable et n'inclut pas d'hypothèses sur la distribution. Par exemple l'épaisseur d'une chape de béton peut être variable à plus ou moins 0,5cm quelle que soit son épaisseur mais le débit d'infiltration vari de 10% et est ainsi d'autant plus incertain qu'il est important.

Pour la loi uniforme un intervalle est choisi dans la plage de variation de la variable considéré. Un tirage uniforme est alors réalisé dans ce nouvel intervalle.

La loi normale quant à elle ne dispose pas de limites pour définir une plage de variation, un intervalle de confiance de 95% est donné pour définir cette plage, présentée en Figure II-6. La valeur attribuée à une variable a alors 95% de chance de se trouver dans l'intervalle de variation. Le reste de la distribution est considérée comme marginal et pourra être coupé pour que les valeurs prises par la variable restent dans le domaine d'étude.

Echantillonnage avec des lois de distribution correspondants aux incertitudes permettant après simulation d'estimer directement la moyenne et la variance attendu pour ces incertitudes (Liu, 2001).

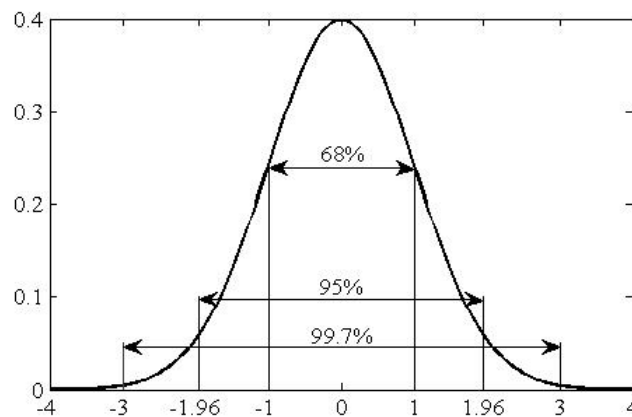


Figure II-6 : Intervalle de confiance défini pour une loi normale $\mu=0$ $\sigma=1$

4.3 Les méthodes d'échantillonnage

La suite de cette partie présente les méthodes de tirage utilisées pour étudier le modèle. Ce sont des processus stochastiques qui sont employés pour générer les points de l'espace des entrées pour l'évaluation du modèle. On différencie stochastique et aléatoire dans le sens où c'est un aléa qui est maîtrisé et contrôlé voire dans certain cas parfaitement défini afin de représenter au mieux les lois de distribution définies.

La première propriété à respecter vis-à-vis des méthodes d'analyse employées est la non-corrélation des entrées. Cette propriété a déjà été évoquée dans la présentation des techniques d'analyse de sensibilité section 2.1. Evaluer l'effet indépendant de chaque entrée requiert de pouvoir distinguer les variations qu'il y a eu sur chacune de ces entrées. Il est donc nécessaire de sélectionner des variables ou de disposer d'un modèle qui a des entrées qui puissent être modifiées de façon indépendante et sans contrainte. Cette propriété à elle seule ne suffit cependant pas à assurer une étude correcte, il est nécessaire de ne pas produire cette corrélation directement dans le choix des valeurs prises par les entrées au cours des essais. Cette condition est obtenue par l'emploi d'un générateur de nombre aléatoire correctement initialisé.

Dans la suite quelques méthodes couramment employées pour améliorer l'efficacité d'un simple générateur aléatoire sont détaillées.

4.3.1 Amélioration de la génération aléatoire classique : Méthode LHS

Pour minimiser le temps de calcul, des techniques ayant besoin d'un nombre réduit de simulation ont été employées. Cependant lors d'un tirage avec des jets aléatoires classiques cette faible quantité d'essai peut rendre difficile l'exploration correcte de l'espace des paramètres où la probabilité d'occurrence est la plus faible.

Une méthode courante de génération des entrées pour les analyses de type Monte-Carlo est la méthode LHS (Latin Hypercube Sampling) (McKay et al., 1979). La loi de distribution est découpée en autant d'espaces équiprobables qu'il y a d'essai. Un tirage uniforme dans chacun de ces sous-espaces permet une représentation correcte de la loi de probabilité quel que soit la taille de l'échantillon. Il est ainsi possible d'obtenir une représentation plus stable des indices malgré une réduction du nombre de simulations (Helton et Davis, 2003). Dans une analyse plus récente H. Janssen discute sur l'avantage d'employer ce type de générateur vis à vis d'autres techniques de génération (Janssen, 2013).

Dans le cas où $N=500$ la distribution est découpée en 500 intervalles de dimension variable, mais disposant chacun d'une probabilité d'occurrence égale de 0,2%. Le choix d'une valeur prise aléatoirement dans chacun de ces 500 intervalles permet une représentation correcte de la loi de distribution initiale.

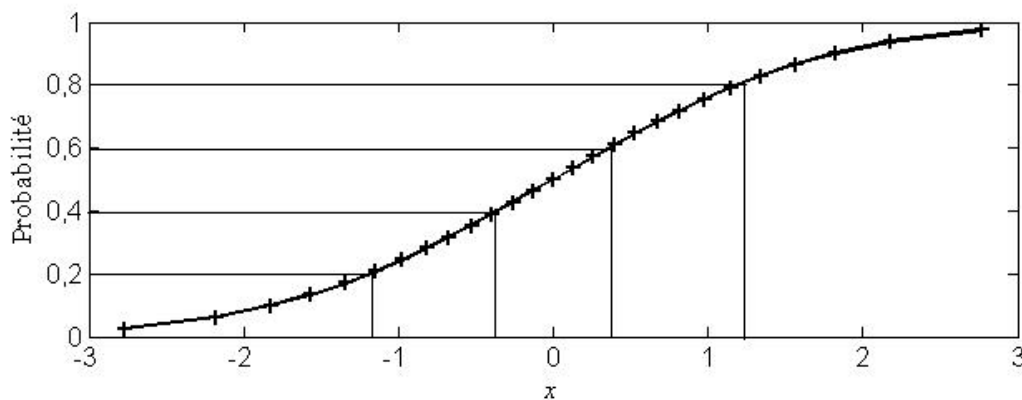


Figure II-7 : Représentation de cinq espaces équiprobables sur une distribution cumulée gaussienne

La méthode se généralise à un espace de m dimensions en réalisant un tirage par variable puis en permutant aléatoirement l'ordre des valeurs de chaque variable. Les échantillons étant générés aléatoirement le nombre de valeurs rend presque impossible la production d'une corrélation suffisante pour perturber les résultats.

Il est néanmoins nécessaire de vérifier la bonne exploration de l'espace qui est maintenant multidimensionnel. Le nombre possible de permutations est très important, on dénombre $(N!)^k$ combinaisons possible et certaines sont plus ou moins optimisées. La Figure II-8 suivante montre la répartition avant et après la permutation aléatoire pour un échantillonnage LHS en dimension deux.

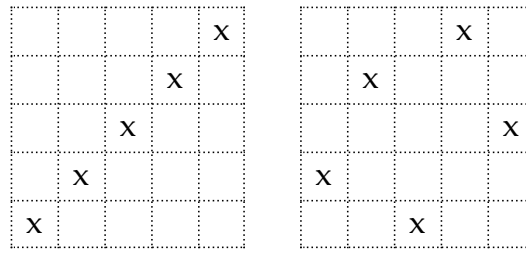


Figure II-8 : Différentes dispersions ordonnées et aléatoire

L'évaluation de la qualité de la répartition est généralement réalisée par une mesure de discrédance ou de distance de type *maximin* ou *minimax* (Johnson et al., 1990).

La discrédance est la capacité à remplir l'espace de façon homogène. La mesure détermine le nombre de points contenus dans un sous-espace par rapport au nombre de points total. Les mesures de distance mesurent la distance de tous les points deux à deux. Pour *maximin* on tend à maximiser la distance minimale entre points. Cette dernière méthode a été choisie pour sélectionner, parmi dix échantillonnages LHS, la répartition qui offrira la meilleure exploration de l'espace.

La principale problématique de cette méthode est que le nombre d'essai est fixé par le découpage initial réalisé sur les distributions. Il n'est alors pas possible d'augmenter progressivement le nombre d'essai. Certains auteurs se sont intéressés à la problématique et proposent des méthodes pour redécouper la distribution en plus d'espaces équiprobables et réaliser des tirages entre les valeurs existantes la méthode prend alors le nom de NLHS pour Nested Latin Hypercube Sampling (Wang, 2003).

Dans ces cas il peut être intéressant de recourir à des suites pseudo aléatoires dont les points sont préétablis et qui fournissent une très bonne exploration de l'espace. Deux de ces suites sont introduites dans la suite.

4.3.2 Suites pseudo aléatoires

Pour générer des points d'évaluation pour le modèle une autre solution couramment employée consiste à utiliser des méthodes quasi-Monte Carlo (QMC). Ce sont des suites présentant une très bonne homogénéité dans l'espace des paramètres, car conçues pour produire une faible discrédance. Elles sont produites à partir d'une logique simple et déterminée, mais qui fournissent des nombres qui semblent aléatoires. L'échantillon produit peut donc être complété en ajoutant les termes suivants des suites considérées.

Deux types de suite ont été implémentés. Les suites de Halton et de Sobol (LP-Tau). Pour illustrer le principe et les limites de ces suites nous allons détailler le principe des suites de Halton (Kuipers et Niederreiter, 2006). Pour plus de détail sur

différentes suites existante le lecteur peut se référer aux travaux de Franco (Franco, 2008).

Les suites de Halton peuvent se voir comme le ratio entre la base des nombres premiers et une base fixe choisie pour chaque variable afin de fournir un nombre entre zéro et un. La suite en base deux commence par les éléments suivants :

$$\frac{1}{2}; \frac{1}{4}; \frac{3}{4}; \frac{1}{8}; \frac{5}{8}; \frac{3}{8}; \frac{7}{8}; \frac{1}{16}; \dots \quad (\text{II.52})$$

En pratique, le sixième terme est construit en retournant l'écriture du nombre six en base deux avec une décimale puis en repassant en base dix.

$$6,0 = (110,0)_2 \mapsto (0,011)_2 = 2^{-2} + 2^{-3} = \frac{3}{8} \quad (\text{II.53})$$

Ces suites sont cependant confrontées à quelques problématiques à haute dimension. Les points générés peuvent en fonction des dimensions observées produire des schémas répétitifs. La Figure II-9 présente les premiers 1000 points des suites de Halton pour les dimensions un et deux puis un cas de schémas répétitifs observé pour les dimensions 11 et 15.

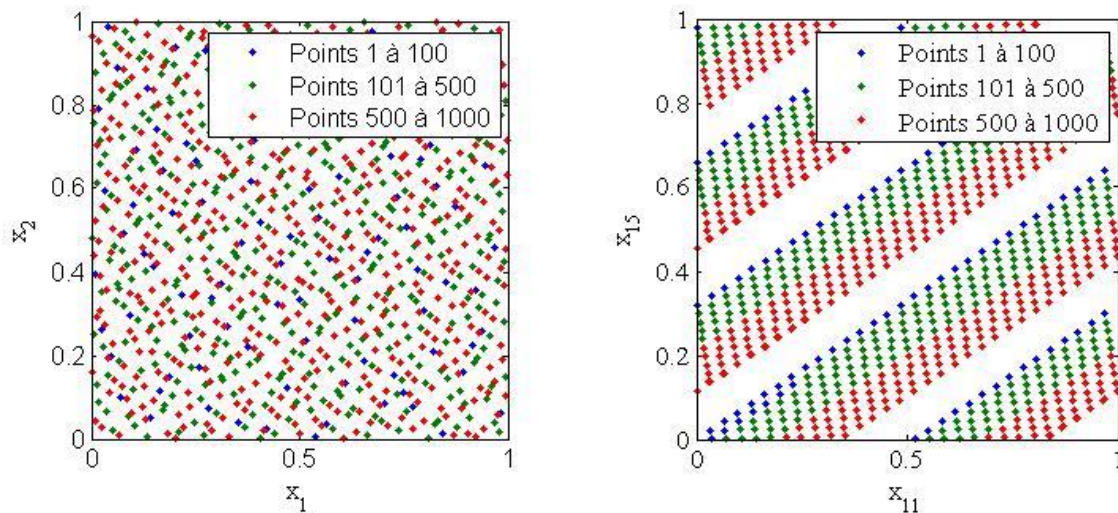


Figure II-9 : Échantillonnage via les suites de Halton (Scatterplot)
(A : Dimensions 1 et 2 ; B : Dimensions 11 et 15)

Une problématique similaire est présente dans les suites d'I.M. Sobol, lesquelles sont construites à partir de polynômes irréductibles particuliers. Le détail d'un algorithme pour la génération de suites LP-Tau est donné par B. Shukhman (Shukhman, 1994). La dispersion des premiers points est très bonne, mais lors de l'augmentation du nombre de point, les points supplémentaires ont tendance à se grouper comme le

montre la Figure II-10. Ce phénomène peut être corrigé en effectuant une initialisation correcte en fonction de la dimension et du nombre de points à fournir. Mais il est difficile de prédire les valeurs d'initialisation qui donneront des répartitions correctes, des tables de valeurs optimales sont donc fournies par quelques auteurs.

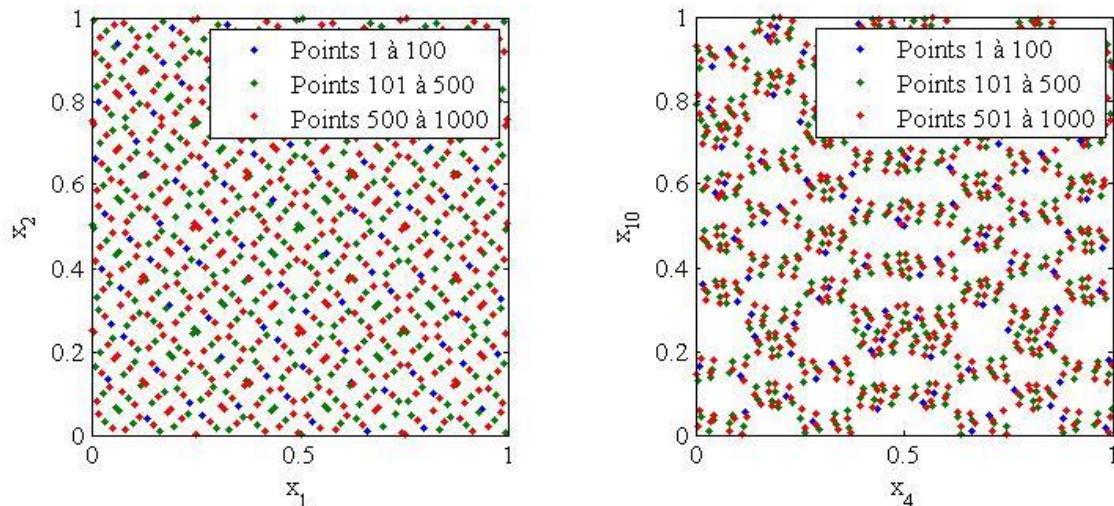


Figure II-10 : Echantillonnage via les suites de Sobol (Scatterplot)
(A : Dimensions 1 et 2 ; B : Dimensions 4 et 10)

Des méthodes dites de Scrambling (Fang et al., 2005) peuvent alors être employées avec ces deux méthodes afin de corriger leurs problématiques.

Il est à noter que ces suites fournissent une exploration selon une loi uniforme de l'espace, pour les utiliser avec des lois de probabilité particulières une fonction de distribution cumulée est utilisée pour modifier la distribution comme présentée pour la technique LHS dans la Figure II-7.

5 Un besoin d'intégration de l'analyse de sensibilité pour l'étude des modèles thermiques

Dans ce chapitre nous avons présenté quelques méthodes applicables en thermique du bâtiment : les méthodes FAST et de Sobol. Le domaine de l'analyse de sensibilité offre de nombreuses autres méthodes pour analyser les modèles. A. Saltelli (Saltelli et Bolado, 1998) indique d'autres techniques capables d'extraire la variance pour évaluer les effets principaux. Iman et Hora (Iman et Hora, 1990) présentent une mesure d'importance (Importance measure), B. Krzykacz et M. McKay (Krzykacz, 1990; McKay, 1995) un ratio de corrélation (Correlation ratio). D'une façon générale plus la quantité d'information demandée est importante plus le nombre de

simulations nécessaire devra être élevé afin d'obtenir une évaluation correcte de l'espace.

Dans le cadre de l'étude des bâtiments qui est envisagé il est nécessaire de mesurer l'importance des entrées de la réhabilitation sur les différentes sorties :

- sur différentes plages de variation des entrées,
- avec le minimum d'hypothèses sur le modèle,
- en intégrant les interactions ou une évaluation globale de celles-ci,
- en réduisant au maximum le nombre de simulations nécessaires.

Les méthodes retenues sont applicables quel que soit le modèle et donc le choix des indicateurs de sortie choisi. Elles permettent d'évaluer les principaux effets du modèle avec une assez bonne confiance et très peu de simulations. Il sera possible d'évaluer cette fiabilité via un estimateur d'erreur.

D'autre part ces méthodes sont applicables quel que soit la plage de variation des entrées et seront donc capables d'évaluer pour les grandes plages de variation la contribution d'une variable sur le résultat. Ce type d'analyse permettra d'établir des pistes d'améliorations en réhabilitation. Pour les variations plus petites et relatives aux incertitudes, ces méthodes vont déterminer quelles sont les variables localement influentes et qui pourront être responsables d'un écart entre simulation et réalité.

Les effets calculés sont enfin clairement interprétables et ne nécessitent pas de connaissances particulières. Pour ce qui est de l'utilisation, la méthode RBD-FAST permet de très rapidement mettre en avant les principales entrées influentes. Elle sera utilisée pour sonder les grandes plages de variation, car son analyse permet de rapidement discerner un comportement sinusoïdal au sein d'un bruit aléatoire. La technique de Sobol est également utilisée, elle offre une meilleure évaluation des entrées peu influentes, mais nécessite un nombre plus important de simulations. Elle sera donc principalement utilisée pour réaliser des analyses de sensibilité sur de faibles plages de variation. Il est à noter que les écarts entre ces méthodes sont relativement faibles et l'analyse de sensibilité est utilisée de façon qualitative pour aider l'utilisateur.

Le chapitre suivant présente dans une première partie le programme qui a été construit afin de lier ces méthodes au logiciel de simulation thermique dynamique et permettre l'exploitation des simulations. Dans une seconde et troisième partie des analyses seront réalisées.

Equation Chapter (Next) Section 1

Chapitre III : INTEGRATION DANS UNE DEMARCHE DE COMPREHENSION DU MODELE

L'analyse de sensibilité apporte les moyens (méthodes et techniques) pour mieux comprendre le comportement des modèles numériques. De nos jours, la réalisation de ce type d'analyse sur les modèles complexes de bâtiments est facilitée par l'augmentation de la puissance de calcul des stations de travail, mais également des outils de simulation dynamique. En effet, les simulations thermiques dynamiques sont de plus en plus employées et permettent de réaliser rapidement des évaluations assez précises pouvant influencer les différents choix de conception. Il apparaît ainsi le besoin d'analyser le comportement du modèle vis-à-vis des valeurs d'entrées choisies afin de valider les choix qui se basent sur ces outils.

L'analyse de sensibilité en thermique du bâtiment a déjà été exploitée dans de nombreuses études : pour la recherche de solution, la calibration, la réhabilitation ou l'étude du parc (Hopfe et Hensen, 2011; Lam et al., 2008; Moran et al., 2012; Sun et Reddy, 2006) ; et semble se généraliser (Tian, 2013). On remarque effectivement qu'autour du contexte de performance énergétique et des méthodes présentées dans le Chapitre II, il y a un réel avantage à mettre en place un outil d'analyse. Mais pour l'instant toutes ces applications ne disposent pas d'une base commune, ce sont des études faites au cas par cas.

Les méthodes d'analyse, dite de sensibilité, employées étudient le modèle via la variation des paramètres d'entrée. Les logiciels de simulation thermique dynamique intègrent généralement des options permettant de faire varier automatiquement les valeurs d'entrée pour réaliser notamment des analyses paramétriques –c'est-à-dire une variation individuelle de chaque entrée. Des outils offrant plus de possibilités

existent et offrent un couplage automatique qui permet de réaliser des variations plus complexes, par exemple jEPlus (Zhang 2009) permet depuis août 2012 de réaliser des échantillonnages type LHS. Ce sont des éléments de première importance qui permettent d'évaluer le comportement du modèle en fonction des configurations. Les capacités de ces méthodes sont cependant limitées à la génération d'échantillonnages simples. Il est alors impossible d'utiliser les techniques évoluées qui ont été présentées dans le Chapitre II.

La communauté de l'analyse de sensibilité a de son côté mis en place un certain nombre d'outils dédiés à faciliter et généraliser la réalisation d'analyses statistiques sur les modèles. Parmi les plus connus on peut citer le projet GNU « R » et « OpenTurn » de EDF. Le groupement de recherche *MASCOT-NUM* (GDR MASCOT-NUM) recense un ensemble d'outil et de routines spécifiques à ces logiciels, ils disposent ainsi de toutes les principales méthodes d'analyse nécessaires pour l'étude des modèles, quels qu'ils soient. Cependant, chaque nouvel utilisateur doit faire face à ce large choix de méthodes et doit procéder à différents couplages pour réaliser les simulations nécessaires.

De plus, la gestion des résultats de simulation est une des limitations majeures de ces outils qui demande une extraction des résultats au cas par cas. Or, la simulation thermique dynamique offre la possibilité d'analyser de nombreux aspects du bâtiment. Il est parfois intéressant de combiner certaines sorties détaillées de la simulation pour avoir des informations particulières en fonction du bâtiment ou du comportement à étudier. Cependant, le modèle peut disposer d'une géométrie complexe avec des dénominations plus ou moins explicites ce qui rend la manipulation des sorties difficile. Il est alors nécessaire de rendre la création d'indicateurs élaborés plus simple et généralisable afin de favoriser l'analyse multicritère. La partie simulation et gestion des résultats est donc mise en défaut et assurer une gestion avancée des sorties semble être une nécessité dans un outil d'analyse de simulation thermique.

C'est pourquoi un outil générique d'analyse de sensibilité pour la simulation thermique dynamique, rapide d'utilisation et le plus automatique possible a été voulu. Pour mettre en place les bases de ce que doit être un outil d'analyse de sensibilité pour l'application à la thermique du bâtiment. Cette partie est destinée à présenter l'outil qui a été mis au point et détaille la méthode employée pour extraire les résultats de la simulation au-delà des informations générales comme la consommation annuelle.

1 Le programme

1.1 Plan général du programme

Les méthodes d'analyse présentées dans le Chapitre II sont dites non intrusives et s'appliquent à des modèles de type boîte noire. Dans ce type de méthode, la liste des opérations devant être réalisées est toujours similaire. Dans une analyse des différentes applications de l'analyse de sensibilité aux systèmes énergétiques complexes dont font partie les modèles de bâtiment W. Tian décrit cette méthodologie en 6 étapes (Tian, 2013) :

- déterminer les variations ;
- créer le modèle énergétique ;
- simuler ;
- collecter les résultats ;
- calculer les effets ;
- présenter les résultats.

La procédure de simulation automatisée reprend donc ces grandes étapes d'une analyse de sensibilité auxquelles sont ajoutées quelques étapes intermédiaires pour les besoins de l'étude. La structure du programme développée est décrite en Figure III-1. La partie du programme dans laquelle apparaît une boucle *méta-modèle* est décrite en section 3 et 4 du Chapitre IV.

Pour l'étude thermique du bâtiment, le logiciel de simulation EnergyPlus a été choisi pour de nombreuses raisons. C'est un outil complet, gratuit et multiplateforme avec un développement important et une capacité de modélisation des systèmes énergétiques développés. Il permet un accès à de nombreuses variables internes du calcul. Les entrées et les sorties sont au format texte et l'outil peut être utilisé en parallèle sans interface graphique. De plus, il a fait l'objet de nombreuses validations et dispose d'un fort soutien de la communauté scientifique (Fumo et al., 2010).

Pour la partie analyse, le code a été développé sur MATLAB pour sa gestion du format matriciel et des séries temporelles. Le langage se veut également accessible et multiplateforme et peut être compilé pour fonctionner de façon autonome. De plus, l'utilisateur de l'outil dispose par l'intermédiaire de MATLAB de nombreux outils statistiques grâce à la Toolbox statistique –non utilisée– ou bien par les routines alimentées par la communauté internationale d'analyse de sensibilité (Joint Research Centre website).

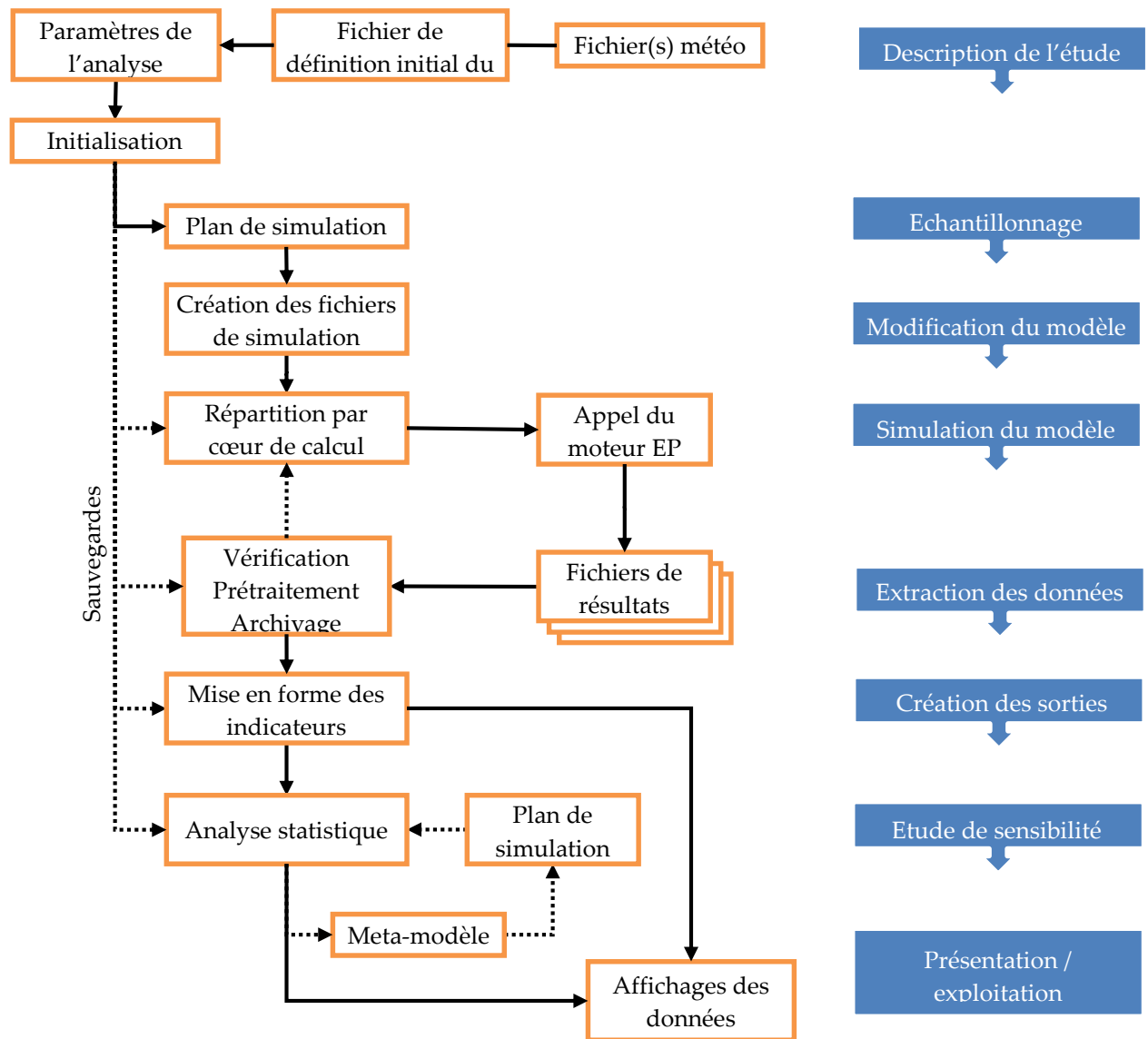


Figure III-1 : Architecture de l'outil

La première étape est la définition des informations nécessaires au bon déroulement de l'analyse. Les principales informations de configuration demandées sont :

- fichier de définition du modèle thermique à analyser,
- variables d'entrées,
- sorties demandées,
- paramètres de l'étude (nombre de simulation, type d'échantillonnage et d'analyse).

D'autres informations de configuration sont présentes, mais ne seront pas explicitées, car secondaires vis-à-vis du principe de fonctionnement.

La Figure III-1 présente également les principaux points de sauvegarde qui sont réalisés régulièrement durant l'avancement du programme afin de reprendre une analyse interrompue ou de réutiliser les résultats pour faire une étude avec une méthode différente.

Dans les sections suivantes, les principales étapes du programme seront détaillées afin de développer le processus de réflexion qui a accompagné le développement de l'outil.

1.2 Définition des entrées et génération des fichiers de simulation

Les simulations thermiques dynamiques sont réalisées avec le logiciel de simulation EnergyPlus. Les nombreuses données d'entrée nécessaires sont généralement regroupées dans un unique fichier de définition qui comprend les données du bâtiment, des systèmes, des scénarios et des options de simulation propres au logiciel thermique. Cette configuration est ensuite associée, pour la simulation, à un fichier météorologique horaire. Dans le déroulement de l'étude, il est nécessaire de réaliser différentes évaluations du modèle dans lesquelles certaines de ces propriétés d'entrée seront modifiées. C'est donc à partir de ces fichiers que différentes variantes de la simulation initiale seront créées. Un fichier de simulation et/ou météo est produit pour chacune de ces variantes, la configuration particulière de chaque fichier est préalablement décidée par l'échantillonnage qui détermine les valeurs prises par les différentes variables considérées pour l'analyse.

Il est d'une part nécessaire de décrire le plan de simulation qui détermine les valeurs à intégrer dans chaque fichier de simulation. Celui-ci est défini en fonction des variables entrées et des propriétés statistiques associées. Puis il s'agit de définir une méthode pour produire à partir de la simulation de base les nouveaux fichiers de simulation qui intègrent les valeurs données par le plan de simulation.

La partie suivante commence par la description de la méthode d'identification dans le fichier initial. La définition qui doit être faite pour les entrées est ensuite explicitée.

1.2.1 Modification du modèle

Les informations nécessaires à la simulation sont regroupées dans les fichiers initiaux de simulation qui se composent de la définition du bâtiment (fichier *.idf) et des sollicitations extérieures (fichier *.epw). Il sera nécessaire d'intervenir sur ces fichiers afin de modifier les valeurs des entrées que l'on souhaite inclure dans l'étude.

Pour la définition des données climatiques, le fichier météo est une entrée particulière de la simulation. Celui-ci contient les valeurs horaires de différentes sollicitations extérieures et de nombreuses corrélations sont présentes dans ces données. Par exemple la température à huit heure est fortement dépendante des températures précédentes, mais également liée aux autres sollicitations telles que le rayonnement solaire ou bien l'humidité. Réaliser une modification simple –c'est-à-dire indépendante entre variables– de ces valeurs n'est donc pas possible sans dégrader l'intégrité du comportement climatique. Cependant la somme des informations contenues dans le fichier météo peut être considérée comme une entrée discrète, par exemple les données de deux stations météo différentes. Il est également possible de coupler cette entrée à des données externes afin d'associer un fichier météo différent pour chaque simulation. Le programme présenté a ainsi pu servir de base au développement d'une méthode de génération sur des données issues de séries temporelles corrélées afin de permettre l'analyse de sensibilité sur des données climatiques (Goffart, 2013).

Pour les données de simulation, une variante du fichier idf initial est créée pour chaque simulation. La réalisation de ces fichiers nécessite de localiser préalablement dans le fichier initial les modifications à apporter. Une méthode courante consiste à remplacer chaque valeur devant être modifiée par une balise spécifique telle que *\$epIsolant\$* dans l'exemple ci-dessous représente l'épaisseur d'isolant :

Material,	
isolation murs,	!- Name
Rough,	!- Roughness
\$epIsolant\$,	!- Thickness {m}
0.0350000,	!- Conductivity {W/m-K}
12.0000000,	!- Density {kg/m3}
840.0000000;	!- Specific Heat {J/kg-K}

Figure III-2 : Extrait du fichier de définition *.idf – localisation directe de la variable

Lors de la création des fichiers de simulation, cette balise est identifiée et remplacée par la valeur correspondante pour la simulation actuelle. Cette méthode est rapide, mais efface la valeur d'origine et rend le fichier initial inexploitable si la balise n'est pas remplacée par une valeur correcte.

La méthode qui est employée consiste à ajouter une ligne de commentaire en amont afin de localiser indirectement la valeur à modifier. La marque de commentaires « ! » du logiciel de simulation est utilisée pour intégrer les informations de localisation dans le fichier initial sans perturber son fonctionnement. Cette localisation indirecte a été choisie pour deux raisons importantes. D'une part, l'intégrité du fichier est

conservée et permet de garder une compatibilité avec tous les outils classiques basés sur le cœur de calcul EnergyPlus – EP-Launch, EP-Editor, DesignBuilder, *etc.* Cette compatibilité semble de première importance lors d’une phase de conception, car en cas de besoin il sera possible de modifier le fichier initial tout en gardant l’identification des variables de réhabilitation. Cette méthode permet d’autre part d’inclure par défaut toutes les identifications nécessaires pour ensuite n’activer que celles qui seront utiles pour l’analyse réalisée. C’est-à-dire choisir au dernier moment, quelles sont les variables à étudier pour la recherche de solution ou l’analyse d’incertitude.

Les entrées qui ont été définies dans le fichier initial de la simulation sont potentiellement modifiables et peuvent donc être prises en compte de différentes façons. Le choix entre trois options permet de modifier rapidement une valeur ou de changer le nombre de variables analysées. La première option définit l’entrée comme une variable, elle est considérée comme une entrée à analyser et va être modifiée à chaque simulation pour étudier son comportement à travers le modèle. La seconde option fixe l’entrée à une valeur personnalisée, l’entrée est considérée comme un paramètre et ne fait plus partie des entrées analysées. Cette option permet de réaliser rapidement des modifications sur certaines entrées sans intervenir sur le fichier initial. La dernière option consiste à désactiver complètement l’entrée pour retourner à la valeur initiale définie dans le fichier.

Afin d’identifier directement cette ligne de commentaire particulière présente avant l’objet que l’on souhaite modifier, la marque de commentaire « !\$ \$ » est utilisée. La ligne est ensuite composée par le nom de la variable ciblée et sa localisation dans les lignes suivantes.

```
!$ $ epIsolant ③
Material,
    isolation murs,      !- Name
    Rough,              !- Roughness
    0.2000000,          !- Thickness {m}
    0.0350000,          !- Conductivity {W/m-K}
    12.0000000,         !- Density {kg/m3}
    840.0000000;        !- Specific Heat {J/kg-K}
```




Figure III-3 : Extrait du fichier de définition *.idf – localisation indirecte de la variable

Dans l’exemple ci-dessus la valeur de la variable *epIsolant* qui correspond à l’épaisseur d’isolant va remplacer la valeur présente à la troisième ligne. Ce type de remplacement est valable pour des données numériques ou textes, ce qui peut

permettre d'inclure différents choix de modèle (exemple : calcul des coefficients convectifs).

La localisation peut également cibler plusieurs lignes pour une même variable en écrivant le numéro de ligne à la manière de MatLab. Par exemple [1:2:9] pour les lignes 1 3 5 7 9 ou {1,[5:7]} pour les lignes 1 5 6 7. Cette définition est particulièrement utile pour la modification de scénarios. Il est aussi possible de réaliser des opérations simples lors du remplacement en ajoutant un opérateur (+,-,*,/).

L'exemple suivant utilise ces deux propriétés afin de moduler la taille du masque solaire dû au débord de toiture en fonction de l'épaisseur d'isolant. La valeur de la variable *epIsolant* définie pour cette simulation est soustraite aux composantes Y des deux premiers vecteurs. Il est ainsi possible de conserver un débord de toiture constant quel que soit l'augmentation de l'isolation extérieure. Et de façon plus générale, ceci permet d'évaluer les multiples impacts que peut avoir la modification de certaines variables.

```
!$$ epIsolant -[6,9]
Shading:Zone:Detailed,
    debord_toiture,          !- Name
    face_sud,                !- Base Surface Name
    ,                        !- Transmittance Schedule Name
    autocalculate,          !- Number of Vertices
    5,                       !- Vertex 1 X-coordinate {m}
    -.4,                     !- Vertex 1 Y-coordinate {m}
    2.5,                     !- Vertex 1 Z-coordinate {m}
    0,                       !- Vertex 2 X-coordinate {m}
    -.4,                     !- Vertex 2 Y-coordinate {m}
    ...
```

Figure III-4 : Extrait du fichier de définition *.idf – opération sur les valeurs présentes

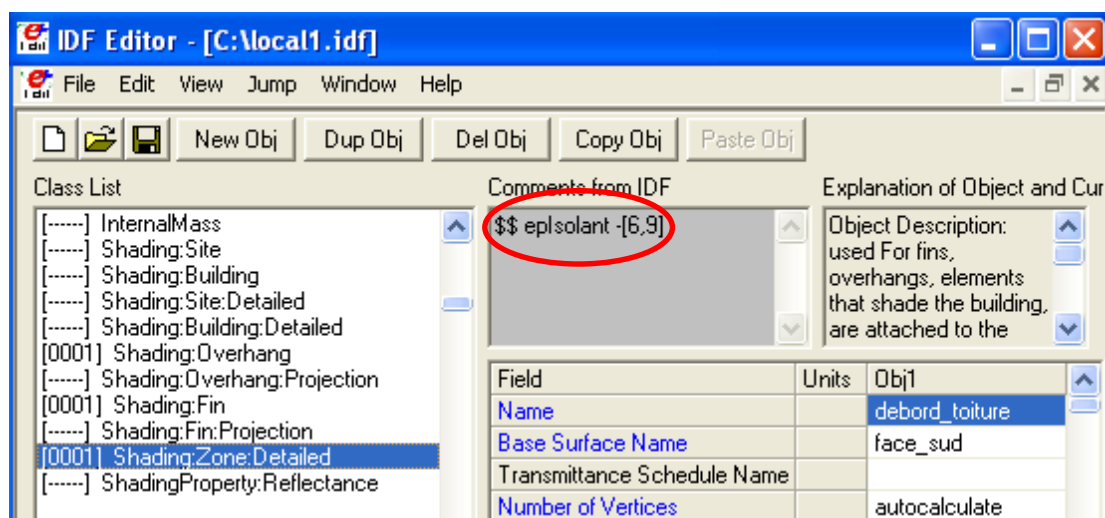


Figure III-5 : Vérification d'un commentaire dans EP-Editor

Cette modification suit parfaitement les règles de construction du fichier IDF et les modifications apportées sont directement visualisables dans le programme d'édition d'Energy Plus nommé EP-Editor comme le montre la Figure III-5.

La méthode de localisation présentée offre ainsi une grande souplesse d'utilisation et permet de réaliser des modifications complexes.

1.2.2 Définition statistique des entrées

Dans le Chapitre II :4.3, les méthodes d'échantillonnage qui ont été définies demandent de déterminer les propriétés statistiques des entrées qui sont analysées. Les propriétés données aux entrées c'est-à-dire les variations produites sur les entrées doivent correspondre au cadre de l'analyse qui est réalisée.

L'objectif est d'une part d'identifier le comportement du modèle, mais également d'identifier les incertitudes dues au choix d'une solution. Il est donc pour l'instant nécessaire de produire au moins deux définitions pour les valeurs d'entrées : L'une pour la recherche de solution qui doit être réalisée avec des plages de variation uniforme sur de larges plages, car il ne peut y avoir d'*a priori* sur la solution ; L'autre avec des variations plus faibles et des lois de probabilité correspondant aux incertitudes relatives à la méconnaissance des propriétés d'entrée de la solution mise en place. Le passage entre les deux définitions est possible avec l'utilisation d'un méta-modèle présenté en Chapitre IV :2.

En principe l'état initial dispose également d'incertitudes non négligeables dues à la méconnaissance du bâtiment et aux méthodes d'évaluation. La calibration du modèle permet de réduire cette incertitude via la comparaison entre mesures et simulations. Il est tout à fait possible de réaliser ce type d'analyse via le programme présenté à

condition d'adapter les entrées étudiées et la définition des données de sortie analysées. La calibration doit ainsi permettre d'ajuster au mieux ces variables de calibration pour ensuite réaliser la recherche de solution. Une fois la solution définie, ces incertitudes peuvent au besoin être réévaluées pour estimer l'erreur faite sur les sorties en considérant la solution et l'état initial. Dans le cas présent le modèle est supposé idéalement calibré et nous ne considérons pas d'incertitudes relatives à l'état initial.

Pour les propriétés statistiques, deux définitions sont donc intégrées : La première détermine la plage de définition de la variable pour la recherche de solution avec des valeurs minimales et maximales pour chaque variable ; La seconde intègre une valeur moyenne par défaut (valeur prise lorsque la variable est fixée), une loi de distribution et son amplitude. Une particularité de la loi de distribution est qu'elle peut être définie de façon fixe ou relative à la moyenne. Par exemple, l'incertitude présente sur l'épaisseur d'isolation peut être de plus ou moins 1cm ou de plus ou moins 10% de la valeur moyenne définie.

Les informations demandées sont donc :

- une variable d'activation qui permet d'activer la variable (1), de la fixer à une valeur particulière (0) ou de la désactiver pour laisser la valeur d'origine (-1),
- un nom d'identification unique pour chaque entrée définie,
- une loi de probabilité (R : relatif),
- une moyenne et une amplitude,
- une plage de variation composée d'une limite haute et basse.

```
1, 'epIsolant', 'GaussianR', [0.20 10], [0.001 0.3];  
1, 'debit_HVAC', 'GaussianR', [0.000305 10], [0.000005  
0.00305];
```

Ces informations permettent ensuite à la méthode d'échantillonnage de générer les valeurs qui seront intégrées dans les fichiers de simulation.

1.3 Réalisation des simulations

Une recherche de solution ou plus généralement l'étude approfondie du modèle nécessite une certaine puissance de calcul. L'utilisation des méthodes précédemment décrites en Chapitre II permet une évaluation précise des effets avec un nombre de simulations restreint. La mise en place de ces méthodes réduit ainsi fortement le coût de simulation et permet de réaliser les calculs demandés sur un ordinateur personnel. Le nombre de simulations est tout de même important et nécessite de

mettre en place des simulations en parallèle afin de solliciter au maximum la puissance de calcul disponible et réduire ainsi le temps de calcul.

Cette parallélisation peut se faire de deux manières. D'une part les ordinateurs actuels disposent généralement de plusieurs cœurs de calcul, la simulation thermique n'est pas un calcul parallèle à proprement dit, mais lancer en parallèle un calcul par cœur permet d'accroître de façon significative la rapidité du calcul. D'autre part, il est possible de répartir la charge de calcul au travers de plusieurs postes en définissant un nombre de cœurs égal à la somme de tous les processeurs utilisés. Il est alors nécessaire de copier les fichiers de simulation sur d'autres postes et ne lancer qu'une partie des lots de simulation.

La répartition par cœur s'effectue en déterminant le groupe de la n-ième simulation par :

$$groupe = 1 + (n - 1) \bmod N_{\text{coeurs}} \quad (\text{III.1})$$

Ces lots de simulation sont définis par des fichiers batch (*.bat) qui appellent chaque simulation du groupe les unes après les autres. Ces appels se font via une version modifiée du fichier d'appel à EnergyPlus afin d'enregistrer les résultats de chaque simulation dans des sous-répertoires distincts. Les simulations restent classiques et peuvent être déportées sur une ou plusieurs stations de calcul à la seule condition de disposer de la bonne version d'EnergyPlus sur le poste.

Les résultats sont vérifiés au cours ou à la fin des simulations afin de s'assurer de la bonne réalisation des calculs. En cas de problème ou d'interruption en cours de simulation, les simulations incorrectes sont identifiées afin de repérer les erreurs probables au niveau des entrées de simulation. Un jeu réduit de simulations est ensuite proposé afin de réitérer le cycle de simulation uniquement sur les essais incorrects ou non réalisés.

Avec la pratique, il s'est avéré que la réalisation d'études sur de larges plages de valeurs d'entrée peut amener à de rares cas où la simulation échoue. Il s'agit de cas limites particulièrement défavorables qui produisent des erreurs systématiques. Cet effet a principalement été observé pour les simulations couplées à un système aérodynamique. Cependant l'échec de quelques simulations sur un large panel d'étude ne faussera pas de façon significative les résultats. L'utilisateur peut choisir d'ignorer les simulations incorrectes et poursuivre avec le jeu de simulation incomplet, s'il le juge suffisant. Par exemple sur mille simulations, l'échec de cinq d'entre elles ne produira pas un biais important. La qualité des résultats n'est donc théoriquement pas

garantie, mais peut tout de même être exploitée. De plus, quel que soit le nombre de simulations correctes une analyse visuelle peut être réalisée.

1.4 Extraction et traitement des sorties

1.4.1 Principe général

Les simulations donnent accès à une quantité importante d'informations sur les échanges du bâtiment, des systèmes et de l'environnement. Les sorties peuvent être demandées avec une définition horaire ou agrégées pour obtenir une valeur annuelle. L'accès à ces sorties ne consomme pas de ressources, mais celles-ci doivent être demandées avant calcul afin d'être intégrées dans les résultats. Il est alors préférable de demander les principales variables sous forme détaillée même si elles ne seront pas utilisées pour ne pas risquer de devoir refaire toutes les simulations. De plus, il n'est généralement pas nécessaire d'étudier l'intégralité des résultats, une sortie peut cibler de nombreux éléments et ceux-ci sont alors tous présents dans les résultats. Par exemple, une température de surface peut contenir les valeurs pour chaque surface quel que soit le nombre de zones.

Il n'est donc pas évident de manipuler toutes ces données à cause de la quantité importante de valeurs et particulièrement quand les surfaces n'ont pas des noms explicites. Mais ces résultats servent à la construction des indicateurs, qui doivent pouvoir analyser des points ou des périodes particuliers. Il a donc été décidé de réaliser une extraction intelligente des sorties qui s'adapte au modèle et facilite la mise en forme des indicateurs. Le traitement des sorties se fait donc en trois étapes : Prétraitement, Extraction et Formatage.

Le prétraitement a pour rôle d'identifier la composition du modèle de bâtiment, de nommer les éléments et d'obtenir leurs caractéristiques. Lors de l'extraction, les variables de sortie présentes sont alors liées à la géométrie et il est possible de formater automatiquement des résultats partiels en fonction des zones, des orientations, de périodes, *etc.* L'utilisateur peut exploiter ces informations afin d'observer différents aspects de la simulation et aller au-delà de la simple étude de la consommation annuelle. L'objectif est de pouvoir :

- construire des indicateurs quel que soit le nombre de zones chauffées ou non
- faire des moyennes de certains facteurs (surfactive ou volumique)
- analyser les défaillances (par élément/ par orientation/ ...)

1.4.2 Analyse du modèle et Extraction

Le prétraitement permet d'extraire les données temporelles et géométriques qui sont communes à toutes les simulations. On suppose pour cela que l'étude est réalisée sur une ou des périodes constantes et le bâtiment conserve sa composition afin que les mêmes sorties soient présentes dans tous les fichiers de résultats. Ce prétraitement fournit des identifiants à toutes les zones et les surfaces, informe sur leurs propriétés et les rassemble par groupes (zones, orientations, parois vitrées/opaque, conditions limites...). On obtient le squelette du bâtiment ce qui permet d'identifier les résultats et de localiser leur origine dans la géométrie quel que soit le modèle.

Ces informations sont extraites d'un des fichiers de résultat qui fournit l'interprétation faite du modèle par le programme de simulation. Le rapport contient entre autres des détails sur toutes les surfaces en incluant le nom, les zones associées, l'aire, les dimensions estimées, l'orientation, l'inclinaison, *etc.* Une base de données est alors produite à partir de ces informations.

D'autre part les résultats de simulation proviennent des fichiers de sorties *.csv mis en forme par EnergyPlus. Les données présentes correspondent aux sorties qui ont été demandées lors de la simulation et sont généralement nommées par « KeyValue: Type (Frequence) » tel que la dénomination :

- « Frequence » représente l'intervalle d'affichage de chaque valeur, allant du pas de temps de calcul à l'année ;
- « Type » correspond à la variable demandée et son unité ;
- « KeyValue » correspond à l'élément auquel s'applique le Type, la valeur * affiche tous les éléments.

Il est nécessaire d'identifier ces données, de les stocker dans un format facilement manipulable. Le système d'extraction utilise pour cela différents opérateurs génériques issus du prétraitement qui prennent lors de l'extraction le nom (%ZONE%, %ZONECONF%, %SURFACE%, %WINDOW%) ou l'identifiant (%ZONEID%, %SURFID%) de tous les éléments qu'ils désignent. L'utilisation de ces identifiants permet une sélection et un traitement des données fiables, puissants et automatiques. L'utilisateur définit uniquement les noms des variables d'EP et les noms des variables locales de la façon suivante :

```

'Environment:Site Outdoor Air Drybulb Temperature
[C] (Daily)', 'donnees.unique.T_ext';
'%ZONE%:Zone Operative Temperature [C] (Hourly)',
'donnees.zone.T_op(:,%ZONEID%);
'%ZONECONF%:Zone Thermal Comfort Fanger Model PMV
[] (Hourly)', 'donnees.zone.PMV(:,%ZONEID%);
'%WINDOW%:Surface Window Heat Loss Energy [J] (Daily)',
'donnees.surface.E_win_loss(:,%SURFID%);
'%SURFACE%:Surface Inside Face Convection Heat Gain Energy
[J] (Hourly)', 'donnees.surface.E_conv(:,%SURFID%);

```

Figure III-6 : Définition des variables de sortie à extraire

Chacune des lignes précédentes correspond à un type de sortie EnergyPlus et est liée à une variable locale. Pour la première ligne, la température extérieure moyenne de chaque jour est enregistrée localement dans la variable *donnees.unique.T_ext* qui forme alors un vecteur de donnée. La seconde ligne correspond à la température opérative de chaque zone à chaque heure et est liée à la variable locale *donnees.zone.T_op* qui forme ici une matrice. Les opérateurs génériques %ZONE% et %ZONEID% permettent dans ce cas d'enregistrer les résultats de chaque zone dans chaque colonne de la variable locale. Le numéro de la colonne correspondant à l'identifiant de l'élément considéré dans la base de données précédemment créée.

Après extraction les fichiers de sorties EnergyPlus peuvent être supprimés. Cela représente une très forte réduction de la masse de données (4% de la taille initiale) et permet une lecture rapide par la suite.

1.4.3 Construction des indicateurs

La dernière étape avant l'analyse est alors la construction des indicateurs de sortie qui permettront de réaliser différentes évaluations du modèle de bâtiment étudié. Les données extraites sont donc traitées une dernière fois pour l'analyse, mais sachant qu'il s'agit de données déjà formatées la lecture des résultats est rapide. Cette partie du programme est la seule qui nécessite une intervention au sein du code, mais est grandement facilitée par la mise en forme des différentes informations faites précédemment. Elle permet de produire rapidement différentes études à condition que les informations de bases soient présentes. Il est par exemple possible de changer la période d'étude ou d'analyser le comportement des parois par orientation.

Une opération courante dans la création des indicateurs est la réalisation d'une moyenne pondérée. Par exemple obtenir la température moyenne opérative sur les zones où une évaluation du confort est réalisée. Il est nécessaire d'identifier dans la base les identifiants des zones disposant d'une évaluation du confort –présence des

occupants. Ces identifiants permettent alors d'une part de sélectionner les bonnes évolutions de température et d'autre part d'obtenir les surfaces ou les volumes des zones correspondantes afin de réaliser une somme pondérée. Une fois toutes ces informations obtenues la pondération peut être réalisée.

De nombreuses mises en forme des sorties sont possibles en suivant le même principe. Pour avoir un aperçu des informations, disponibles on peut citer pour les zones :

- Nom
- Volume
- Surface au sol
- Surface en contact avec l'extérieur
- Surface vitrée extérieure
- Nombre de surface composant la zone
- Théorie(s) de confort appliquée
- Nombre de personnes
- Points de mesure de la carte d'éclairement
- Présence d'infiltrations
- Présence d'un système de ventilation

Et pour les surfaces :

- Nom
- Type (murs / vitrage / encadrement de fenêtre / porte)
- Zone de référence
- Surface d'appuis
- Orientation
- Inclinaison
- Condition limite opposée
- Surface nette
- Surface totale

A partir de là différentes listes de parois sont extraites pour obtenir directement des groupes par orientation ou par type. Le groupe Nord associé au groupe Vitrage permet ainsi d'obtenir les informations désirées quel que soit le modèle.

D'autre part, la gestion matricielle des résultats permet une mise en forme rapide et aisée avec la possibilité de réaliser des moyennes horaires ou journalières des résultats avec une simple réorganisation des valeurs.

Une mise en forme particulière des données permet également de réaliser une analyse de sensibilité pour chaque pas de temps de la simulation –ou d’une période définie. Ce principe a été utilisé pour réaliser une analyse de sensibilité afin d’évaluer l’influence des variables au cours du temps. Ce type d’analyse ouvre ainsi une nouvelle voie pour la calibration des modèles.

L’utilisation de ce programme est retranscrite à travers les deux applications suivantes.

2 Exemple d’utilisation : Etude d’éclairement sur un modèle simple de bâtiment

Dans cette première application (Rabouille et al., 2014), le programme est utilisé pour faire varier quelques entrées d’un modèle simple. Aucun indice de sensibilité ne sera véritablement estimé, mais les résultats des simulations sont représentés sous forme de nuage de points appelé *Scatterplot*. Chaque point représente une simulation dont la position est déterminée par la valeur du résultat de simulation –la sortie– sur l’axe des ordonnées et la valeur de l’entrée considérée sur l’axe des abscisses. Ce type de représentation, bien que simple, est couramment employé en analyse de sensibilité pour évaluer qualitativement la corrélation entre une entrée et une ou deux sorties. La forme du nuage permet alors une estimation visuelle du comportement –c’est-à-dire les variations, la linéarité et la dispersion– entre les variables choisies pour chaque axe.

L’objectif de cette analyse est de présenter le programme tout en analysant un aspect intéressant du bâtiment qui peut avoir des influences variables. Il a été choisi d’évaluer dans quelle mesure une isolation extérieure, les vitrages et la protection solaire peuvent influencer sur l’éclairement.

Un modèle simple et un nombre réduit de variables sont pris en compte afin de limiter les sources d’erreur et d’analyser seulement l’impact de la réhabilitation sur l’éclairage naturel. On dispose ainsi d’une plus grande maîtrise sur l’évaluation des résultats.

2.1 Définition du cas d’étude

L’objet de cette première étude est un local présenté en Figure III-7. Celui-ci se veut être le plus simple possible, de forme parallépipédique rectangle il a une surface au sol de 5x2,5 m et d’une hauteur sous plafond de 2,5 m. La structure est constituée de 20 cm de béton lourd avec une isolation extérieure. Il dispose d’une unique fenêtre orientée Sud de 2x1,2 m avec un débord de mur distant de 10 cm et d’épaisseur égale

à l'épaisseur d'isolant. Nous avons également placé une casquette solaire à 30 cm au-dessus de la fenêtre qui dépasse de la surface extérieure du mur de 40 cm. Les parois verticales disposent en plus d'un revêtement intérieur composé d'une lame d'air et d'une plaque de plâtre. Le système de chauffage est idéal, purement convectif et réversible afin d'assurer un refroidissement au besoin. Les consignes de chauffage et de refroidissement correspondantes sont respectivement de 20°C et 26°C. Nous supposons enfin que le local est soumis au climat de Lyon. Toutes les autres variables sont laissées par défaut vis-à-vis du logiciel EnergyPlus.

Ce local simple est destiné à réaliser une première application du programme présenté et permet de mettre en avant certains comportements couplés présents dans le bâtiment. Dans l'étude, l'éclairage intérieur est analysé vis-à-vis de l'effet d'une isolation extérieure, du type de vitrage et de la taille de la casquette solaire. Cet éclairage est une source de confort visuel et thermique, mais également d'inconfort avec le risque de surchauffe ou d'éblouissement.

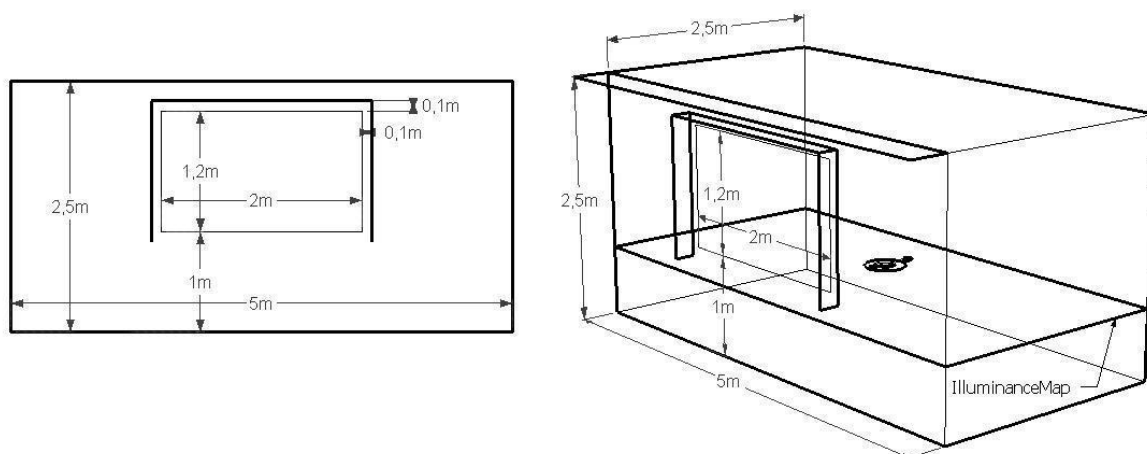


Figure III-7 : Géométrie du local (face sud et vue trois-quarts sud)

Dans le cadre de l'étude, l'éclairage est défini comme l'éclairage moyen sur les heures d'ensoleillement et sur l'ensemble de la surface à une hauteur de 70 cm ce qui est représenté en Figure III-7 par *IlluminanceMap*. Les besoins représentent l'énergie fournie par un système idéal en kWh/m². Sont également présents les transferts (conductifs et radiatifs) des vitrages vers la zone thermique en kWh/m² de vitrage. Les résultats seront analysés sur les mois de janvier et de juillet.

Un jeu réduit de variables d'entrées est pris en compte et les valeurs associées aux entrées sont définies dans la Tableau III-1. Dans cet exemple nous illustrons l'effet de modifications fortes des entrées pour la réhabilitation. Ainsi l'isolation varie de 0,1 cm à 30 cm. Des variations uniformes seront réalisées dans ces intervalles d'étude avec un échantillonnage LHS.

Tableau III-1 : Variables de la simulation pour l'étude de l'éclairage

Elément	Unité	Minimum	Maximum
Isolants	[cm]	0,1	30
Casquette	[cm]	1	100
Renouvellement d'air	[vol./h]	0,06	2
Vitrage	[6] [6/16/6]	[6(PE)/16/6]	[6/16/3/16/6]

Pour le type de vitrage il a été choisi de définir la variable comme une entrée discrète à 4 niveaux. La définition des vitrages peut être complexe, en plus du nombre de vitres les différentes épaisseurs peuvent varier, disposer de traitements pour réduire l'émissivité et les espaces d'air peuvent être remplacés par un gaz rare type argon. Les propriétés des vitrages sont ainsi très variables en fonction des modèles et des constructeurs et ont donc des comportements complexes qui ne peuvent se résumer à une modification de la conductivité de ceux-ci. De plus, certains vitrages sont conçus afin de répondre à des contraintes spécifiques de type acoustique ou anti-effraction. Il sera donc difficile de produire une liste exhaustive et de comparer les vitrages qui ont des propriétés particulières. La liste des vitrages employés a été composée à partir des matériaux d'EnergyPlus afin de produire quelques assemblages types. L'objectif est de visualiser le rôle du vitrage dans la construction en considérant les principaux types : simple, double, double peu émissif et triple vitrage.

2.2 Analyse du comportement

Pour réaliser une analyse globale du comportement, les valeurs des entrées sont distribuées sur tout leur domaine de définition. Dans un premier temps, l'analyse brute des sorties permet de mettre en avant quelques effets couplés. Etant donné que l'éclairage est le point d'intérêt principal, la Figure III-8 présente sur deux graphiques l'énergie transmise par les vitrages vers l'ambiance en fonction de l'éclairage moyen pour le mois d'hiver et le mois d'été. Les différents types de vitrages sont représentés par différentes couleurs.

Quel que soit la saison l'augmentation du nombre de vitres produit une élévation du flux transmis et réduit l'éclairage intérieur. Le comportement des vitrages est cependant différent en fonction de la saison. En hiver, le type de vitrage a une forte influence et forme 3 groupes distincts. La forte augmentation du flux transmis à l'ambiance intérieure par les vitrages permet dans les conditions hivernales d'améliorer les capacités thermiques du bâtiment. Il est à noter que le double PE et triple vitrage ont des comportements très proches en hiver et sont les seuls qui peuvent avoir un bilan énergétique positif au niveau de la surface. En été, les

différences entre vitrages sont plus faibles, mais quand même visibles. La technologie du vitrage produit principalement une modification sur l'éclairage, l'énergie transmise est très peu impactée et dépend d'une ou plusieurs autres variables qui impactent également sur l'éclairage.

D'autre part, une certaine dispersion est présente dans les résultats et notamment durant l'hiver, cela est dû à l'effet de la variation des autres entrées. Il est alors nécessaire de réaliser une vérification de toutes les autres courbes pour identifier quelle variable est responsable de cette dispersion.

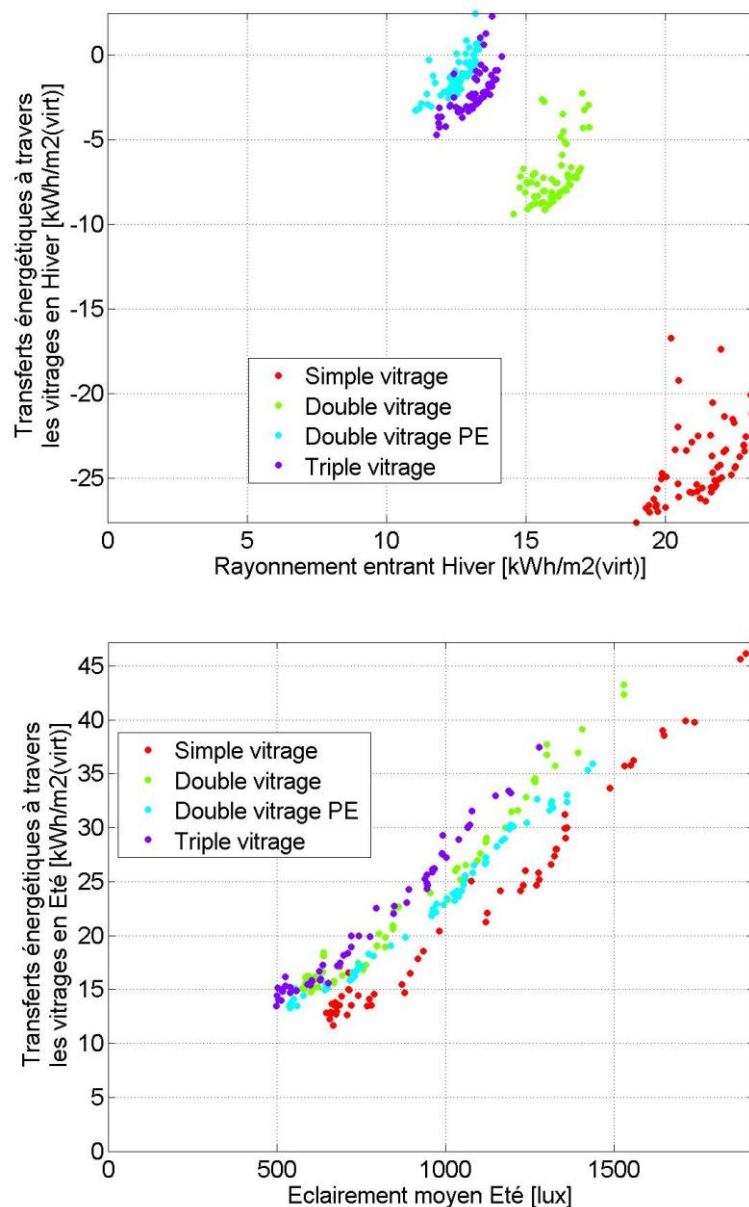


Figure III-8 : Eclairage moyen de la zone en fonction des apports des surfaces vitrées dans la zone – pour les mois de janvier et de juillet –

On note deux variables influentes qui sont logiquement la taille de la casquette en face sud et l'épaisseur d'isolant qui modifie la taille des débords de murs au niveau des fenêtres. Les Figure III-9 et Figure III-10 présentent l'évolution des transferts par les vitrages en fonction de la saison et selon des valeurs croissantes de ces deux variables. Il est intéressant de noter que ces deux variables n'influent pas de la même façon ni aux mêmes périodes. Il est connu que le débord de toiture influe de façon notable en été quand le soleil est haut alors que les débords de fenêtre sont plus à même de masquer les rayonnements horizontaux en début et fin de journée quel que soit la saison. Ce comportement se retrouve dans les différents graphiques. Par leurs dimensions les débords de mur ont une influence faible, mais qui n'est pas à négliger surtout si le débord de toiture est très réduit ou très en hauteur par rapport aux fenêtres.

Les graphiques permettent de mettre en avant les comportements intéressants et également leurs amplitudes. Il est ainsi possible de faire des choix argumentés à partir de cette analyse.

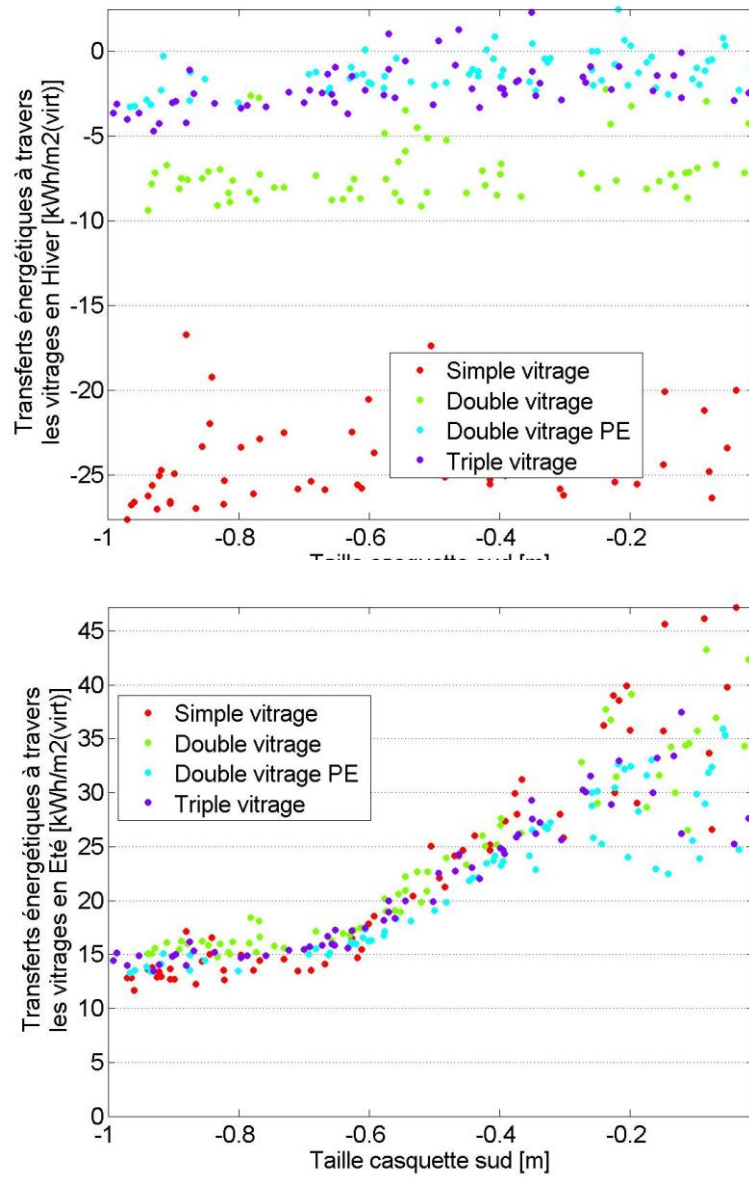


Figure III-9 : Effets de la casquette solaires sur l'énergie entrant par les vitrages

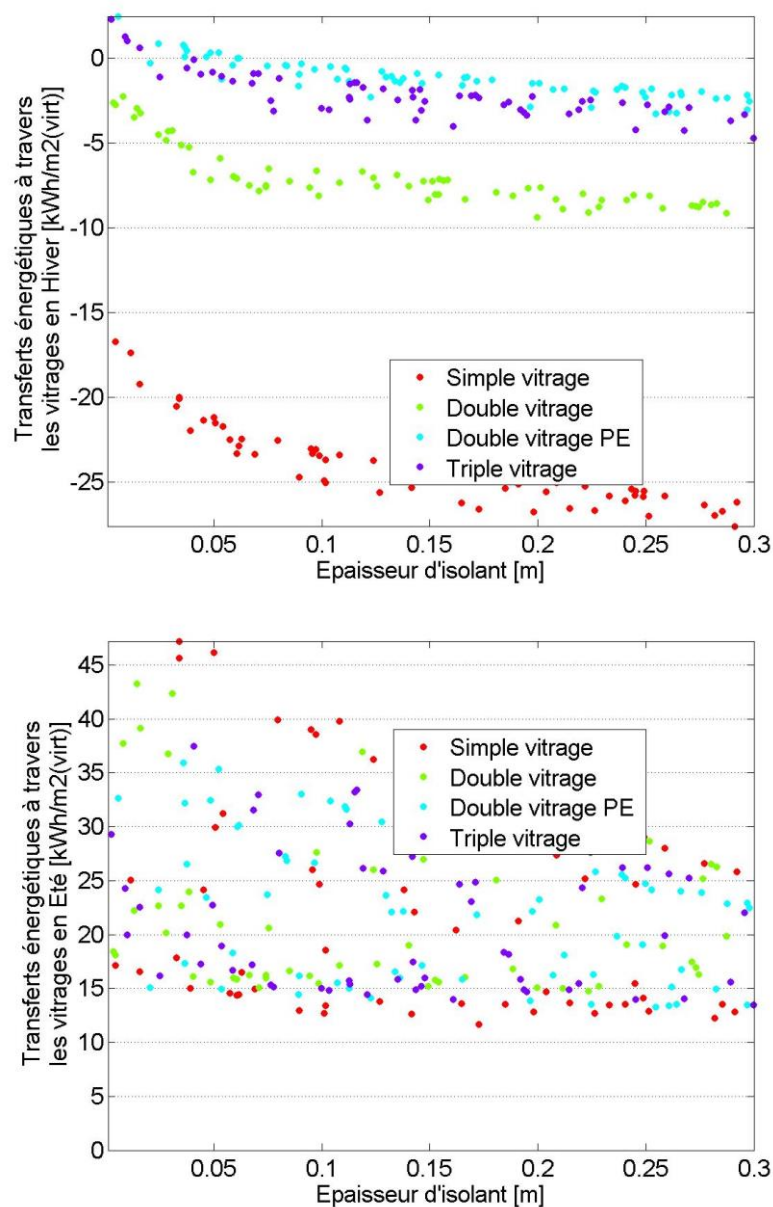


Figure III-10 : Effets des débords de murs sur l'énergie entrant par les vitrages

2.3 Calcul de l'incertitude

L'étude précédente ne permet cependant pas de voir avec précision le comportement local. Pour cela une analyse d'incertitude peut être réalisée sur l'état initial dégradé et l'état final réhabilité. Le Tableau III-2 montre les caractéristiques médianes pour ces deux états. Toutes les variables disposent d'incertitudes sauf le vitrage qui est fixe.

Deux autres analyses sont alors réalisées avec ces nouvelles définitions. L'échantillonnage de ces états est réalisé avec la méthode type LHS présentée en Chapitre II :4.3.1 où deux tirages de 500 points sont réalisés afin d'obtenir une estimation fiable de la variance. Le Tableau III-3 présente les résultats obtenus pour

les besoins, l'éclairage et l'inconfort en été et en hiver. Les valeurs présentées correspondent aux moyennes et entre parenthèse aux écarts types de ces sorties. L'écart type a été choisi afin d'obtenir une valeur de même dimension que la moyenne pour représenter la variance. Afin d'uniformiser la présentation des résultats, ce formalisme est utilisé également même si la dispersion des résultats n'est pas gaussienne.

Tableau III-2 : Valeurs d'entrée prises en compte pour l'état initial et final
– incertitude Gaussienne : 95% de confiance soit $1,96\sigma$ –

Elément	Unité	Moyenne état Initial	Moyenne état Final	Incertitudes
Isolants	[cm]	5	20	Gauss. $\pm 10\%$
Casquette	[cm]	10	60	Gauss. $\pm 5\%$
Renouvellement d'air	[vol./h]	1	0,15	Gauss. $\pm 15\%$
Vitrage	[-]	Simple	Double (PE)	–Fixe–

On remarque que la réhabilitation produit une forte modification du comportement. L'augmentation des performances est due à un meilleur contrôle de tous les échanges thermiques. Les besoins de chauffage passent de 30,3 à 3,22 kWh/m².mois avec une amélioration du confort et également une légère réduction des besoins de climatisation de 8,90 à 7,03 kWh/m².mois.

Mis à part pour l'éclairage, l'opération réalisée permet une augmentation de la confiance pouvant être attribuée sur les besoins. C'est-à-dire que vis-à-vis des incertitudes définies la réhabilitation permet d'obtenir une solution plus stable. La variance entre l'état initial et l'état réhabilité est réduite ce qui indique une plus faible dispersion des résultats malgré des incertitudes plus importantes sur la majorité des entrées.

Tableau III-3 : Valeurs de sortie calculées pour l'état initial et final
– chaque sortie est représentée par la moyenne et l'écart type des résultats –

		Etat initial		Etat final	
		Moyenne	Ecart Type	Moyenne	Ecart Type
Besoin [kWh/m ² .mois]	Hiver	30,3	(0,52)	3,22	(0,054)
	Eté	-8,90	(0,01)	-7,03	(0,003)
Eclairage [lux]	Hiver	966	(0,02)	715	(0,40)
	Eté	1890	(0,37)	733	(124)
Inconfort [°C.jours/mois]	Hiver	78,8	(0,67)	36,9	(1,09)
	Eté	3,73	(0,004)	1,55	(0,003)

2.4 Bilan de l'utilisation du programme

Cette première application permet de mettre en avant deux éléments forts de l'analyse

Certaines entrées ont de multiples effets sur le bâtiment, via leur comportement sur différentes sorties, mais également via les possibles couplages entre ces entrées qui peuvent être inclus dans le programme grâce aux fonctionnalités exprimées en section 1.2. Le programme montre la capacité à gérer de façon simple les entrées du modèle thermique et également à analyser les différentes sorties.

L'étude des Figure III-8 à Figure III-10 a montré qu'en fonction des sorties observées, il est possible d'obtenir des comportements très variables et généralement non linéaires, mais qui restent tout de même monotones. On remarque également que le comportement du bâtiment est différent en fonction des saisons.

Au final, pour la recherche de performance, il est nécessaire d'identifier manuellement quels sont les graphiques dans lesquels une tendance est présente. Dans le cadre de la première application, un nombre très limité de variable et de sortie a été pris en compte, mais pour les cas non simplifiés, le nombre de variables de sortie devient trop important pour permettre la recherche manuelle. D'autre part, pour l'étude d'incertitude, il est important de quantifier des écarts en estimant les variables responsables de l'amélioration ainsi que celles responsables de la dispersion résiduelle. Pour répondre à ces problématiques, dans la partie suivante, l'analyse des performances est poussée à l'estimation du comportement relatif de chaque variable selon différentes saisons grâce à une des méthodes d'analyse de sensibilité précédemment décrite.

3 Evaluation multicritère du comportement d'un bâtiment réel

Pour cette seconde analyse (Rabouille et al., 2013), il a été décidé de se rapprocher d'un bâtiment réel et donc proche de la complexité d'un cas d'étude courant. Pour cela les maisons INCAS –présentées ci-dessous– ont été choisies. Ces maisons conservent une géométrie simple et disposent déjà d'un modèle énergétique validé.

On analyse uniquement le fonctionnement thermique du bâtiment via les besoins, les échanges par les surfaces opaques et le confort été/hiver. L'analyse de sensibilité est utilisée pour déterminer rapidement les entrées influentes de la réhabilitation vis-à-vis des sorties. La méthode RBD-FAST est employée à cette fin. Dans la suite, les résultats des indices de sensibilité obtenus permettent d'identifier les graphiques intéressants pour déterminer les valeurs idéales des entrées.

3.1 Définition du cas d'étude

Le modèle de bâtiment utilisé est celui des maisons INCAS présentes sur le site de l'INES (Institut National de l'Energie Solaire) au Bourget-du-Lac. Ces maisons font parties d'une plateforme expérimentale conçue pour valider les simulations thermiques dynamiques par comparaison entre mesures et simulation. Ces logements passifs sont ainsi bien définis et disposent d'un environnement contrôlé qui offre un cadre particulièrement favorable pour la modélisation thermique tout en conservant la complexité d'un bâtiment réel. Quatre bâtiments de ce type sont présents : Béton banché ; Double mur ; Brique Mono-mur ; Ossature bois.



Figure III-11: Maisons INCAS (image : www.ines-solaire.org)

Le fichier de base considéré est le modèle thermique EnergyPlus de la maison type Double Mur. La ventilation est en simple flux, aucun scénario d'occupation n'est présent et un chauffage idéal convectif assure le maintien de la consigne de température. Ce modèle énergétique est issu des travaux de thèse de Clara Spitz (Spitz et al., 2012), il a été calibré via un processus de validation entre mesure et simulation (Goffart, 2013; Merheb, 2013; Spitz, 2012).

Pour réaliser la comparaison avec un bâtiment à réhabiliter, les propriétés thermiques sont dégradées. Le bâtiment simulé est ainsi constitué d'une structure en parpaing et d'une épaisseur de laine de verre extérieure. Le fichier météo TMY2 de Lyon est utilisé.

Les sorties qui sont étudiées portent sur le confort et la capacité du bâtiment à maintenir ces conditions de confort. Le premier est évalué, quel que soit la saison par la théorie du confort adaptatif pour fournir des degrés jours d'inconfort au-delà de la classe I (écart de 2°C sur la température optimale soit 90% des occupants satisfait). En hiver, les conditions intérieures dépendent des besoins et sont définies par l'énergie totale apportée à l'ambiance intérieure durant le mois d'hiver exprimée en kWh/m².mois. En été, le bâtiment n'est pas conditionné par un système de régulation de la température, il est dit en évolution libre et les besoins sont donc nuls. La ventilation est la principale source d'échange, mais est directement dépendante du

débit choisi et apporte ainsi peu d'information. Il est donc nécessaire de définir une autre sortie pour évaluer globalement les échanges énergétiques du bâtiment. Pour cela, l'énergie arrivant dans l'ambiance par les surfaces opaques est choisie et est exprimée en kWh/m²(surface).mois. Les échanges par les parois sont faibles et cette sortie informe donc sur la capacité du bâtiment à rester proche des conditions extérieures.

Pour évaluer l'impact possible des différentes solutions de réhabilitation, les variables présentées dans le Tableau III-4 suivant sont analysées. Le tableau présente les limites de variation de chaque variable.

Tableau III-4 : Définition des variables

Elément	Unité	Minimum	Maximum
Isolation des murs	[cm]	1	30
Isolation du Vide Sanitaire	[cm]	1	30
Isolation des combles	[cm]	1	30
Renouvellement d'air (infiltrations / ventilation)	[vol/h]	0,1	2
Epaisseur des chapes (pour l'inertie)	[cm]	1	20
Renouvellement d'air (sanitaire et des combles)	[vol/h]	0,1	4
Consigne de chauffage	[°C]	18	21
Technologie de Vitrages	Simple / Double / Double PE		

Toutes les entrées sont échantillonnées avec la méthode LHS selon des lois uniformes afin de rechercher les solutions comme il a été décrit précédemment. Les limites de variations sont définies pour correspondre aux valeurs critiques minimales et maximales qu'il est possible de rencontrer dans un logement. Ces grandes plages de variation permettront comme précédemment d'identifier le rôle des variables vis-à-vis de la performance atteinte puisque tout le domaine d'étude est sondé.

Les cinq premières variables sont des variables courantes de la réhabilitation. Trois autres variables d'entrée sont ajoutées pour vérifier l'influence d'autres phénomènes secondaires. La température de consigne a notamment un rôle important dans la régulation des systèmes et a une influence sur la consommation et les conditions de confort.

Une variable représentative de la qualité des ponts thermique a également été envisagée, mais non incluse. Les ponts thermiques sont des éléments importants en réhabilitation, cependant ces défauts de conception produisent des déperditions qui sont difficilement évaluables. Généralement des estimations en régime statique sont réalisées et des abaques fournissent les valeurs pour les principaux modes

constructifs directement sous forme de déperditions supplémentaires via des coefficients Ψ en $W/K.m$ et X en W/K . Cette définition permet une estimation rapide, mais est en désaccord avec la physique du modèle. Au niveau de la simulation thermique dynamique, EnergyPlus n'a pour l'instant pas les moyens de les évaluer, mais il est cependant possible de les prendre en compte via la création de sous surfaces composées de matériaux aux comportements dynamiques équivalents (ASHRAE, 2001; Nagata, 2005). Cette définition est plus précise, mais demande une modélisation particulière pour chaque point singulier. Quelle que soit la méthode d'évaluation employée, ces défauts sont dus à la structure et ne peuvent que difficilement être traités sauf par une isolation extérieure. Il a donc été décidé de ne pas les inclure dans les variables étudiées, mais de les considérer comme des constantes et de réaliser une isolation extérieure.

Le renouvellement d'air et la ventilation ont été regroupés dans une même variable. D'un point de vue thermique, sans système double flux, ces deux variables déterminent toutes deux la quantité d'air extérieure entrante. D'un point de vue pratique, elles correspondent à des phénomènes bien différents. La ventilation est généralement mécanique et permet de déterminer un renouvellement d'air, généralement via le débit d'air extrait. Les infiltrations sont des défauts de l'enveloppe qui permettent un apport d'air neuf en plus des entrées d'air prévues pour la ventilation, mais produisent également en fonction des écarts de température et de pression, des débits supplémentaires. L'infiltration peut ainsi jouer un rôle sanitaire important en cas de ventilation mécanique insuffisante, cependant l'estimation de ces débits est une tâche complexe qui est pour l'heure difficilement maîtrisée. Les infiltrations peuvent ainsi être vues comme l'élément qui produit des incertitudes vis-à-vis des débits imposés.

Au niveau des débits, le débit de ventilation minimal réglementaire imposé (Arrêté du 24 mars 1982 relatif à l'aération des logements) pour une ventilation mécanique dans un logement de type 4 est de $90m^3/h$, ce qui correspond approximativement à $25m^3/h.per.$ D'un autre côté, la réglementation thermique donne des valeurs pour l'étanchéité à l'air sous $4Pa$. Cet écart de pression de $4Pa$ est retenu pour permettre de donner une valeur de débit qui corresponde à une situation classique. Des valeurs équivalentes de débit d'infiltration et de taux de renouvellement sont présentées dans le Tableau III-5 calculées pour les données du bâtiment.

Ces débits d'infiltration sont ainsi bien supérieurs aux débits imposés pour la ventilation et l'évaluation du renouvellement d'air est ainsi très incertaine.

Tableau III-5 : Valeurs équivalentes de débit d'infiltration réglementaire
(surface extérieure déperditive de 198.7 m² ; volume chauffé de 238.9 m³)

	m³/h.m²	m³/h	Vol/h
RT2005 (pour calcul)	$I_4 = 1,3$	258	1.1
RT2005 (référence)	$I_4 = 0,8$	158	0,67
RT2012 (mesure)	$Q_{4Pa-Surf} = 0,6$	119	0,5

Les valeurs d'étanchéité peuvent cependant être largement améliorées, les labels passifs demandent par exemple de mesurer un débit de 0,6 vol/h avec un écart de pression de 50Pa soit approximativement 0,11 vol/h à 4Pa.

Ces deux variables ont donc été regroupées, la valeur la plus haute a été fixée arbitrairement à 2 vol/h et la valeur basse à 0,1 vol/h pour correspondre à un régime de fonctionnement réduit dans un logement particulièrement étanche. La valeur de débit couvre ainsi toute la gamme de renouvellement d'air allant d'un bâtiment labélisé passif à un logement fortement perméable.

3.2 Etude des indices de sensibilité

L'objectif de l'étude est d'analyser le comportement des différentes variables à partir de 250 simulations. Ce nombre réduit de simulations ne permet pas une exploration parfaite de l'espace, c'est-à-dire que toutes les combinaisons ne seront pas évaluées. Cependant, la méthode d'échantillonnage sélectionne les valeurs d'entrée de la simulation qui permettent une exploration de l'espace uniforme et utilise l'intégralité des domaines de variation des variables d'entrée –voir Chapitre II :4.2 et Chapitre III :1.2.2.

Le comportement général des sorties est évalué vis-à-vis de chaque entrée il peut être facilement visualisé par les *Scatterplot* comme précédemment. On dispose dans cette application d'une plus grande complexité, il y a 8 entrées et 4 sorties soit 32 comportements différents. Le nombre de comportements à analyser peut ainsi augmenter rapidement avec l'ajout d'entrées ou de sorties. L'analyse de sensibilité va permettre d'estimer les entrées les plus influentes sur les différentes sorties et par conséquent d'identifier rapidement les relations intéressantes parmi toutes ces combinaisons. Pour cette analyse, la méthode RBD-FAST est employée.

3.2.1 Première évaluation

Les résultats de l'analyse de sensibilité sont présentés dans le Tableau III-6. Les indices de sensibilité sont exprimés en pourcentage, ils quantifient l'influence des entrées pour chaque sortie. Une valeur élevée indique une forte influence et

réciroquement une valeur faible indique une influence faible. Les valeurs les plus hautes sont mises en gras et sont analysées par la suite.

La valeur entre parenthèses correspond à l'écart type mesuré par Bootstrap. Cette valeur peut être vue comme une représentation de l'erreur commise sur l'évaluation de l'indice de sensibilité.

La dernière ligne contient la somme de tous les effets de chacune des quatre sorties présentées. Selon le principe de l'analyse de sensibilité si cette somme se rapproche de l'unité les interactions contenues dans le modèle sont faibles. Cependant, une marge d'erreur de dix pour cent est généralement tolérée sur la somme des indices. La valeur proche de 80% pour l'inconfort d'été dénote ainsi la présence de quelques interactions. L'analyse ne sera cependant pas plus développée, car pour les besoins actuels il n'est pas nécessaire d'étudier l'origine de cette interaction.

Le tableau permet de présenter de façon synthétique les principales relations entre les entrées et les sorties. Ce premier jeu de résultat montre une différence significative dans le rôle des entrées, les effets sont variables en fonction des sorties observées et d'une façon générale seule quelques entrées dominant. Les variations sur les sorties et donc les choix à réaliser seront donc principalement dépendantes de ces variables.

Tableau III-6 : Indices de sensibilité principaux selon RBD-FAST avec l'écart type de bootstrap

	Eté		Hiver	
	Inconfort été	Parois opaques	Inconfort hiver	Besoins
Epaisseur d'isolant Murs	4,04% ($\pm 2\%$)	2,71% ($\pm 2\%$)	10,4% ($\pm 5\%$)	34,8% ($\pm 5\%$)
Ep. d'isolant Vide San.	6,98% ($\pm 3\%$)	0,14% ($\pm 3\%$)	-1,20% ($\pm 2\%$)	0,12% ($\pm 3\%$)
Ep. d'isolant Combles	5,52% ($\pm 3\%$)	4,72% ($\pm 3\%$)	-2,96% ($\pm 2\%$)	3,99% ($\pm 3\%$)
Inertie	1,08% ($\pm 3\%$)	-2,29% ($\pm 2\%$)	-0,64% ($\pm 3\%$)	-0,91% ($\pm 2\%$)
Renouvellement d'air	58,1% ($\pm 5\%$)	88,3% ($\pm 1\%$)	-1,43% ($\pm 2\%$)	63,5% ($\pm 4\%$)
Ventilation z. non-hab.	1,35% ($\pm 2\%$)	0,14% ($\pm 2\%$)	2,14% ($\pm 3\%$)	-0,08% ($\pm 3\%$)
Technologie de vitrage	2,89% ($\pm 3\%$)	6,54% ($\pm 4\%$)	1,83% ($\pm 2\%$)	-0,31% ($\pm 3\%$)
Consigne de chauffage	-0,31% ($\pm 2\%$)	4,59% ($\pm 2\%$)	81,2% ($\pm 2\%$)	4,55% ($\pm 3\%$)
Somme	79,7%	106%	89,3%	106%

Les graphiques suivants présentent les résultats de chaque sortie en fonction des deux entrées les plus influentes. Les résultats de simulation sont organisés selon les valeurs croissantes de la variable choisie, mais toutes les autres entrées sont

aléatoires. Etant donné que l'ensemble de l'espace est criblé, il est possible d'obtenir pour chaque sortie une estimation des capacités minimales et maximales et également le comportement des entrées les plus influentes. Cette analyse visuelle est qualitative, elle permet d'identifier facilement une tendance et en déduire la plage de variation optimale.

La première sortie présentée porte sur le dépassement du critère de confort en été, lequel est dominé par le renouvellement d'air (58%). La Figure III-12 présente une diminution de l'inconfort sous forme exponentielle négative avec l'augmentation du débit. Ainsi, il n'apparaît pas de réelle problématique de surchauffe si le logement est un minimum ventilé. On remarque au contraire à partir de 1 vol/h une valeur très basse qui implique une augmentation de l'inconfort froid c'est-à-dire une température intérieure qui est inférieure à la température opérative.

La seconde variable influente est l'épaisseur d'isolant du vide sanitaire (7%). Le sol est une source de rafraîchissement, une augmentation de l'isolation limite les échanges avec cette source et produit alors une élévation de l'inconfort. Cet effet est cependant faible et la tendance est difficilement évaluable. L'épaisseur d'isolant du vide sanitaire ne permettra pas de contrôler efficacement la sensation de confort en été. Mais, une ventilation de 0,6 vol/h permet clairement d'obtenir un inconfort chaud sous les 30°C.jours, soit en moyenne 1°C au-dessus de la limite de confort de classe I.

Une autre sortie est présentée pour comprendre un peu mieux le fonctionnement d'été. Pour cela, l'énergie entrante dans l'ambiance par les parois opaque est évaluée Figure III-13. Cette sortie permet de vérifier si le local est plus ou moins en équilibre vis-à-vis de la température extérieure sans perturber l'interprétation par la théorie de confort.

L'augmentation du renouvellement d'air (88%) produit ici une augmentation de l'énergie entrante dans l'ambiance par les parois. La température intérieure est donc en moyenne plus basse que la température extérieure. L'augmentation du débit permet donc d'augmenter les transferts de la paroi de l'extérieur vers l'intérieur et donc de réduire la surchauffe, mais comme nous l'avons vu précédemment ceci peut induire de l'inconfort par des températures basses. En cas de besoin, une régulation de la sur-ventilation nocturne permet de mieux valoriser ce phénomène pour réduire l'élévation de température sans produire de refroidissement trop important. La mise en place de ce type de système peut être évaluée en définissant les critères de mise en route et d'arrêt parmi les entrées.

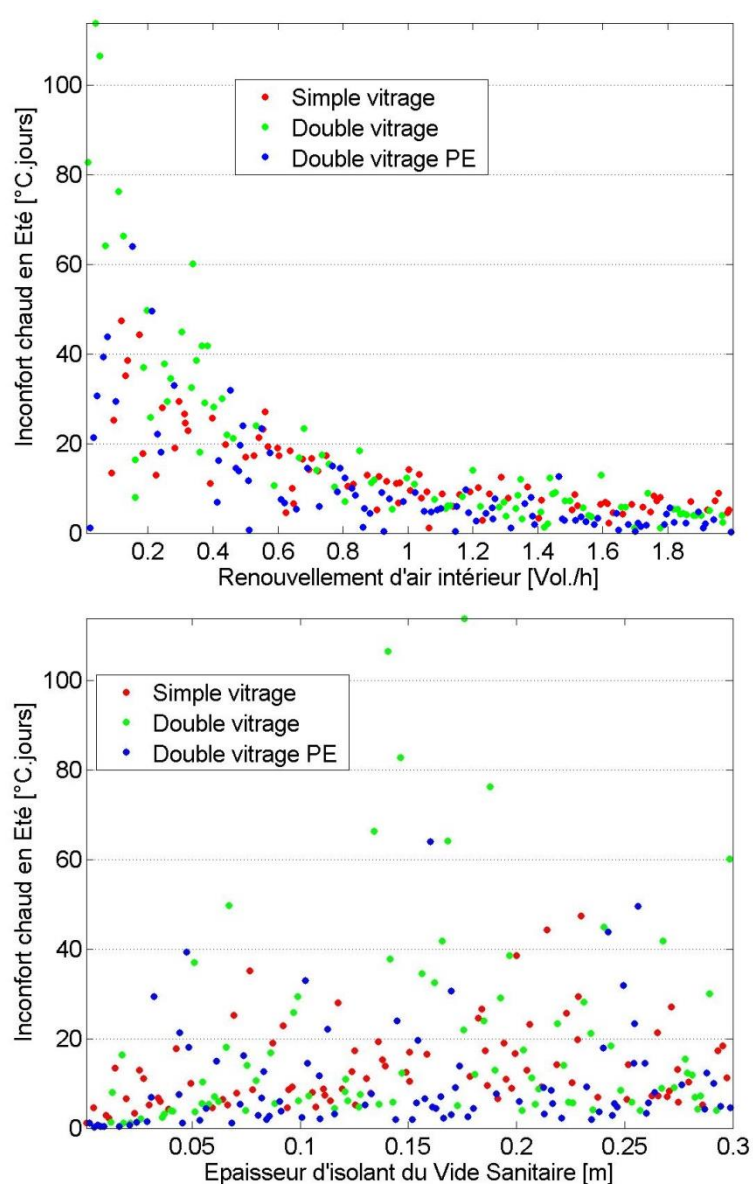


Figure III-12 : Dépassement de la limite haute de confort en été

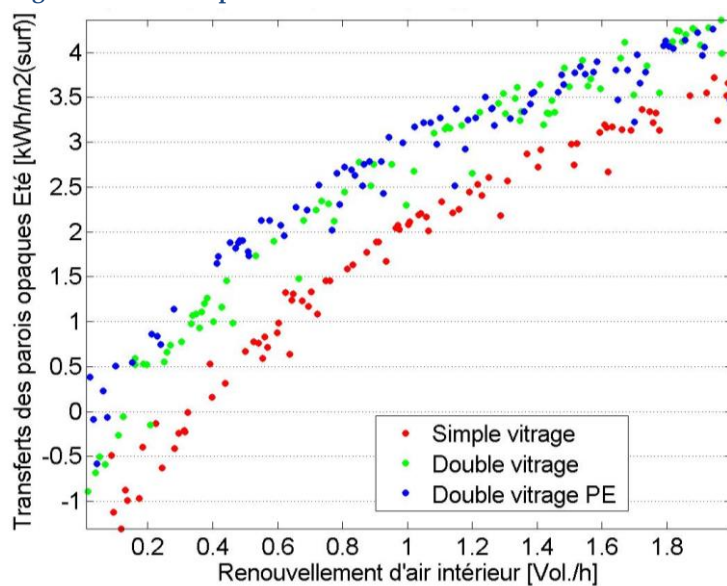


Figure III-13 : Energie entrante dans l'ambiance par les parois opaques

Le type de vitrage apparaît comme la seconde entrée influente (6%) et sa contribution est directement visualisable sur le graphique précédent par la coloration réalisée. Comme dans l'application précédente –section 2.2– les écarts sont relativement faibles et le double vitrage (en vert) est très proche du double vitrage peu émissif (en bleu).

Les autres résultats portent sur le comportement en hiver (Figure III-14 et Figure III-15), durant lequel le bâtiment est contrôlé par le système de chauffage. La puissance thermique est alors adaptée afin que la température d'air atteigne la valeur de consigne fixée. En fonction de la sortie observée, le comportement vis-à-vis des entrées sera différent.

L'évaluation du confort dépend de la température opérative, les entrées importantes sont donc la consigne de chauffage (81%) et l'isolation des murs (10%) qui impacte sur la température de surface. La Figure III-14 représente l'évolution de l'inconfort froid en fonction de la consigne de chauffage d'une part et en fonction de l'épaisseur d'isolant d'autre part. On observe que l'augmentation de l'isolation réduit l'inconfort, mais l'évolution est cependant difficilement visible à cause de la dispersion importante produite par la modification de la température de consigne. Les valeurs les plus basses de degrés jours d'inconfort ne sont cependant atteignables qu'avec une certaine épaisseur d'isolation. En effet, les valeurs en dessous des 60°C.jours ne sont possibles qu'à partir d'épaisseurs d'isolant approchant les 15 cm. De plus à 21°C la consigne ne permet toujours pas d'atteindre le critère de confort. Selon la théorie adaptative, cela est dû au taux de satisfaction demandé élevé qui est atteint pour une température opérative de 20,9°C et à la température de surface intérieure toujours inférieure à la consigne en raison du sens du flux.

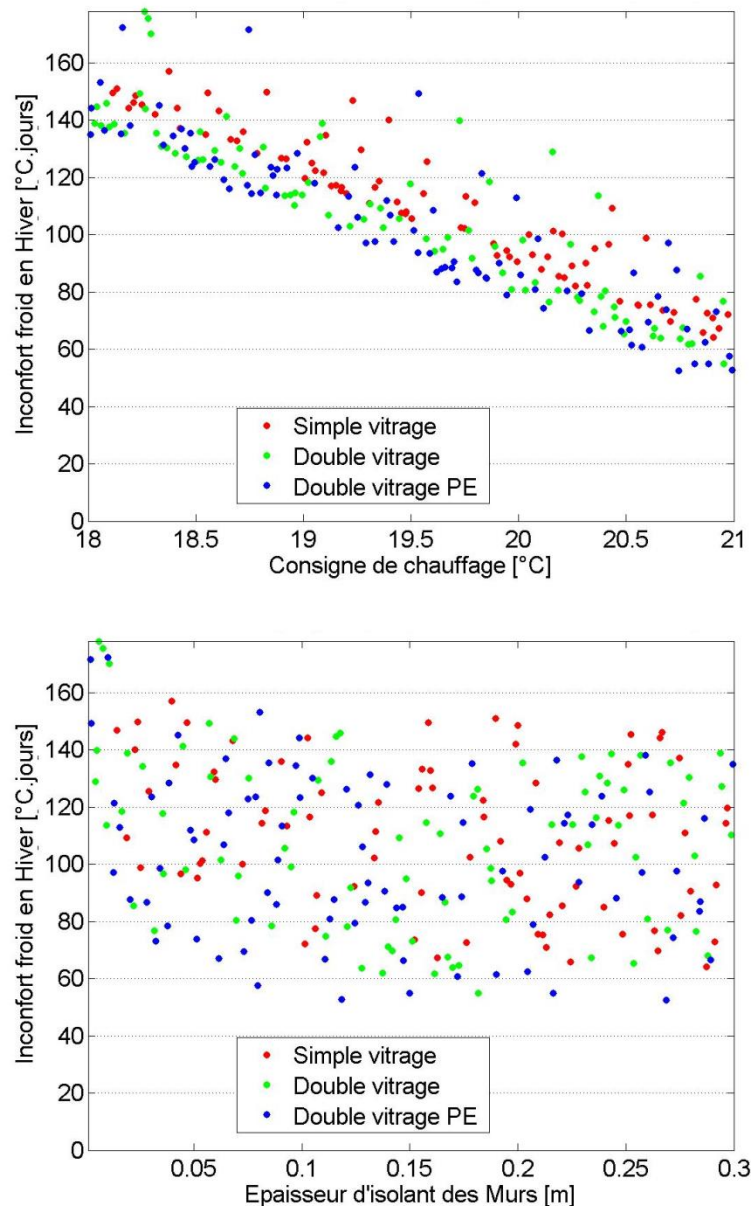


Figure III-14 : Dépassement de la limite basse de confort en hiver

La dernière sortie observée en Figure III-15 correspond au besoin de chauffage nécessaire pour atteindre la température de consigne. Un résultat intéressant est la faible participation de la température de consigne à ce résultat, plus de précision sur ce phénomène est apporté en fin de section. On retrouve principalement les principaux modes de déperdition de chaleur, à savoir le renouvellement d'air (63%) et l'isolation des parois verticales (35%).

L'isolation est le second facteur et une tendance en exponentielle peut être aperçue mais est perturbée par l'effet du renouvellement d'air. Cette tendance sera confirmée en Chapitre IV :4.2 en Figure IV-12 lors de l'utilisation du méta-modèle. On observe en Figure III-15 qu'une isolation de plus de 10 cm semble inutile si aucun travail n'a

été réalisé sur l'étanchéité et les vitrages. En effet, en fonction des valeurs prises par l'étanchéité et le type de vitrage, les besoins de chauffage peuvent varier du simple au septuple ($[5-35 \text{ KWh/m}^2]$).

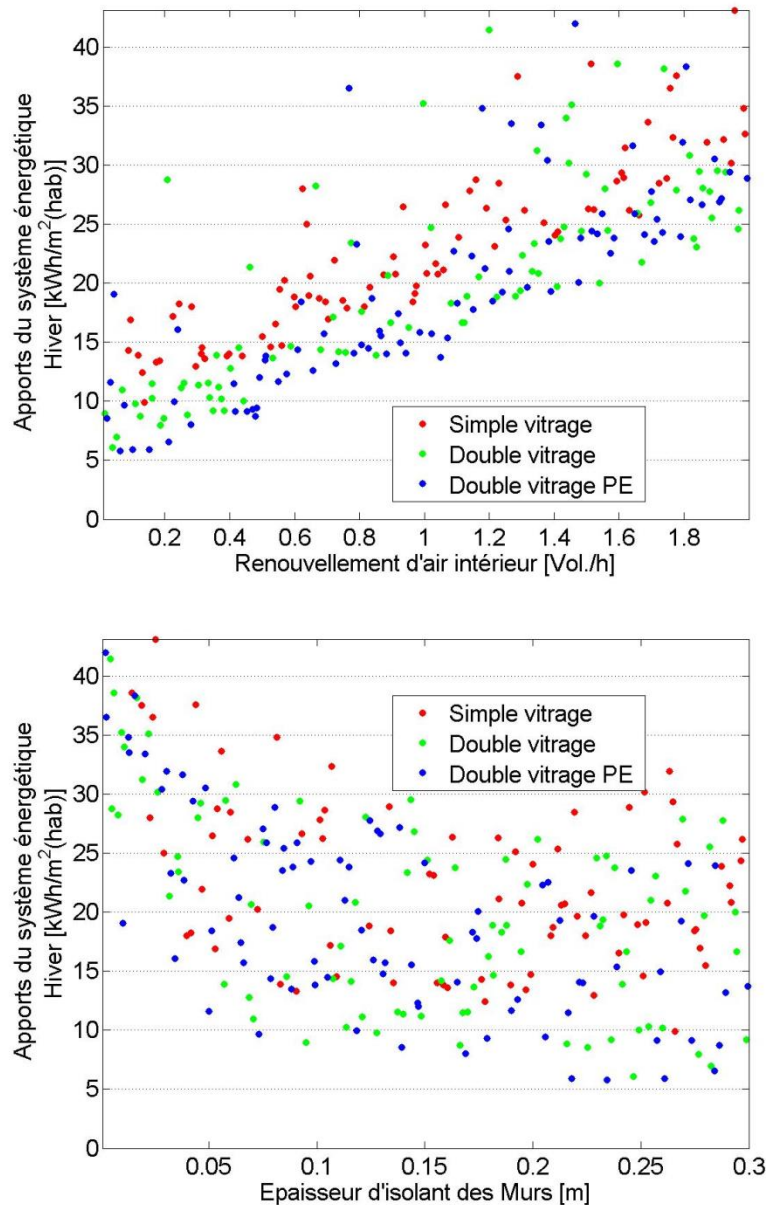


Figure III-15 : Evolutions des besoins de chauffage en hiver

Afin de vérifier la non corrélation entre la consigne de chauffage et les besoins énergétiques, la Figure III-16 suivante montre le graphique correspondant à la relation entre besoin du système et inconfort. Aucune orientation particulière entre ces sorties n'est visible, la majorité des points forment un disque. Cependant ce résultat ne dit pas que la consommation du système est indépendante de la consigne, mais que vis-à-vis des variations présentes cet effet est négligeable par rapport aux effets des autres variables. Les indices de sensibilité et les informations obtenues

dépendent ainsi fortement de la variabilité attribuée aux entrées. Nous verrons en section 3.3 lors de l'évaluation de la solution retenue que des effets différents seront observés et que la consigne sera alors fortement influente en hiver.

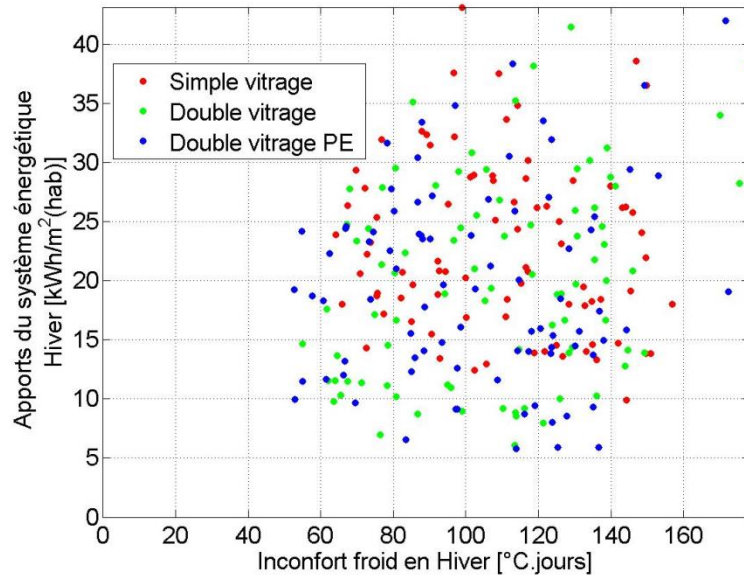


Figure III-16 : Corrélation entre les besoins de chauffage et le confort

Le comportement des variables les plus influentes vis-à-vis des principales sorties ayant été analysé, on remarque que toutes les entrées ont des effets différents au sein de la simulation. Cela nous permet de fixer dès à présent trois variables dans le cas étudié :

La ventilation est la principale variable à maîtriser, elle a un effet linéaire et permet de réduire l'élévation de température en été, mais est également une source de déperdition en hiver. C'est une variable très incertaine une valeur arbitraire de 0,6 vol/h est choisie qui correspond à un effort important au niveau de l'étanchéité. En pratique l'objectif est de réduire au maximum l'infiltration pour gérer au mieux des débits sanitaires et la surchauffe avec des mécanismes de ventilation mécanique contrôlée et parfois aussi naturelle.

L'isolation des murs est une variable principale des besoins, mais son efficacité décroît rapidement, car de forme asymptotique. Une épaisseur de 20cm permet d'avoir une paroi très bien isolée. La valeur définie pour le renouvellement d'air informe cependant qu'il n'est plus possible de réduire les besoins de chauffage sous la barre des 10kWh/m² pour le mois de janvier.

La consigne de chauffage fixée à 20,5°C, principale variable du confort, permet d'ajuster le confort entre 60 et 80 °C.jours d'inconfort pour le mois de janvier.

Tous ces choix restent cependant approximatifs, car ils sont réalisés graphiquement mais ils sont effectués avec une bonne connaissance des capacités possibles et des conséquences sur les différentes sorties. Une fois les principales entrées fixées le comportement du bâtiment est presque connu quel que soient les valeurs d'entrées des autres variables, car la dispersion des résultats est fortement réduite. Une nouvelle évaluation avec les variables restantes permet d'affiner l'analyse.

3.2.2 Seconde évaluation

Un jeu de simulation supplémentaire est maintenant réalisé pour connaître de façon plus précise le comportement des autres variables (dites secondaires) dans le modèle. Les options du programme sont identiques, seules les trois entrées devant être fixées ne sont plus définies comme des variables, mais comme des paramètres –valeurs fixes. Le tableau suivant représente les résultats mis à jour.

Tableau III-7 : Indices de sensibilité principaux selon RBD-FAST avec l'écart type de bootstrap

	Inconfort été	Parois opaques	Inconfort hiver	Besoins
Ep. d'isolant Vide San.	57,0% (4%)	84,4% (3%)	15,8% (5%)	10,8% (5%)
Ep. d'isolant Combles	4,18% (3%)	3,44% (3%)	-1,76% (2%)	-1,96% (2%)
Inertie	15,3% (5%)	-0,85% (3%)	-2,15% (2%)	-2,06% (2%)
Ventilation z. non-hab.	2,79% (4%)	0,05% (3%)	0,65% (3%)	0,73% (3%)
Technologie de vitrage	11,0% (4%)	8,06% (4%)	83,1% (2%)	87,1% (1%)
Total	90,2%	95,1%	95,6%	94,6%

D'une façon générale, on remarque que l'épaisseur d'isolant du vide sanitaire n'ayant pas été fixée, celle-ci reste dans les entrées influentes pour le confort, mais apparaît également sur les autres sorties. La variable présente un indice élevé principalement en été et qui produit un comportement clair sur les graphiques, comme on peut le voir sur la Figure III-17. Le type de vitrage apparaît ensuite comme une entrée dont le comportement domine en hiver, comme le présente la Figure III-18. L'inertie du bâtiment apparaît ensuite avec un effet de plus faible ampleur limité au confort d'été.

Ces nouveaux résultats permettent de réduire encore les écarts présents. La Figure III-18 permet de voir que la variabilité résiduelle devient très faible une fois que le type de vitrage et l'épaisseur d'isolant du vide sanitaire sont fixés. Initialement les besoins variaient entre 5 et 43 kWh/m².mois (Figure III-15), ils sont ensuite réduits entre 12 et 21 (Figure III-18), pour maintenant être à la décimale près autour de 12 kWh/m².

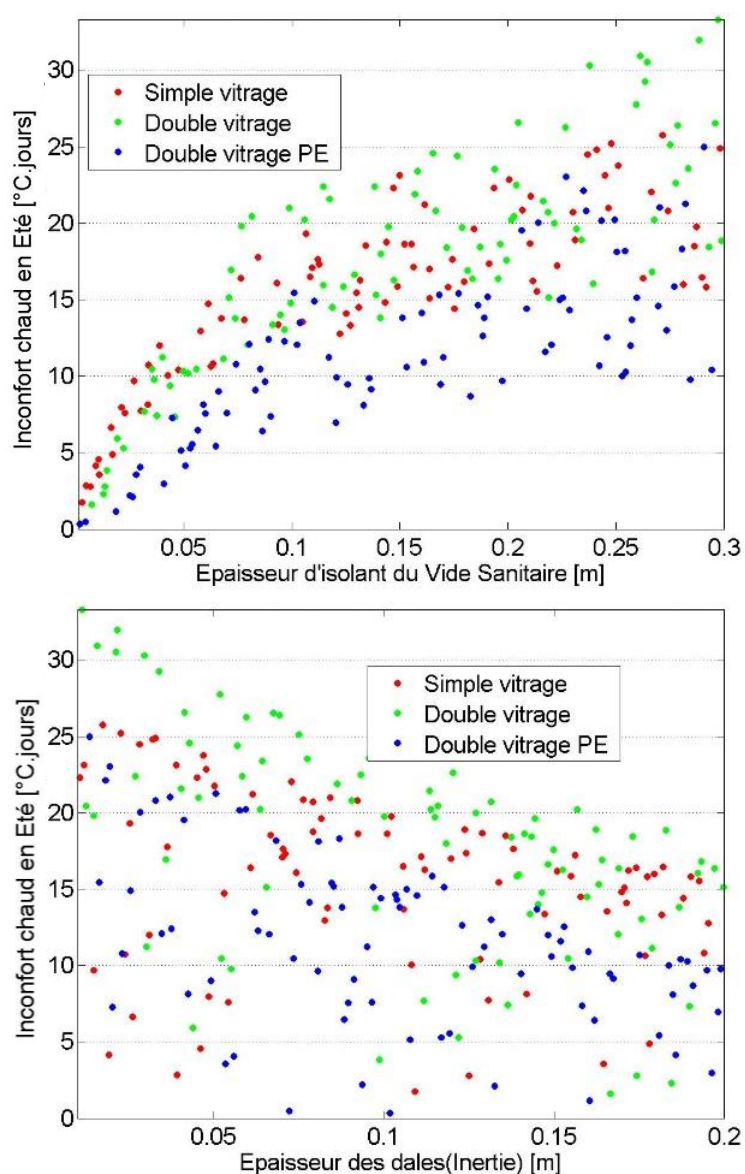


Figure III-17 : Nouvelles évolutions du dépassement de la limite haute de confort en été

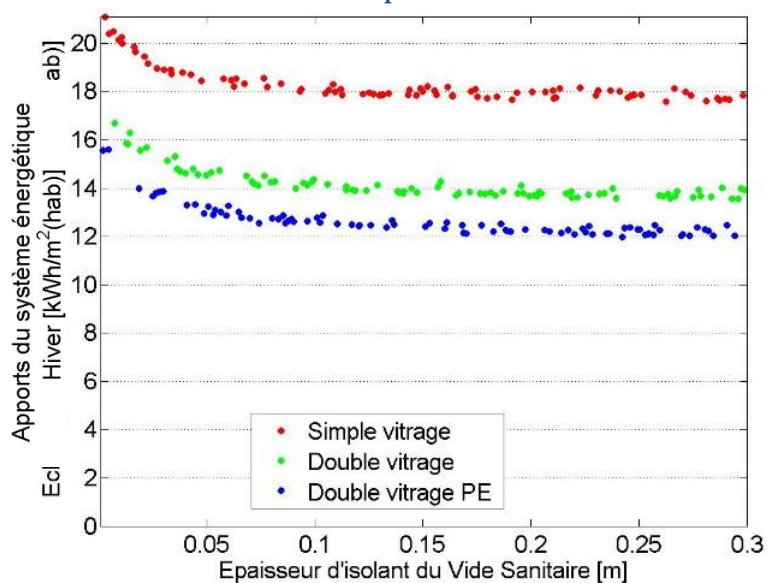


Figure III-18 : Nouvelle évolution des besoins de chauffage en hiver

3.2.3 Classement des variables

Les deux études précédentes ont montré que les indices calculés présentent des valeurs élevées et chaque sortie a un comportement clair vis-à-vis d'une voire deux entrées. La démarche prise jusqu'à présent consiste à fixer les variables les plus influentes afin de réduire la variance après chaque étude.

Dans cette partie, nous nous intéressons à classer les variables par ordre d'influence de façon indépendante pour chaque sortie et sans considérer les interactions. Le Tableau III-8 présente colonne par colonne le classement des entrées réalisé pour chaque sortie. A chaque analyse, l'entrée la plus influente est fixée et une nouvelle analyse est réalisée avec les variables restantes. Les valeurs prises par les entrées fixées orientent l'analyse autour de la solution proposée dans le Tableau III-9.

Entre parenthèses est affiché l'indice de la variable fixée. Un indice élevé indique que la variable avait une forte contribution à la variance. Fixer cette variable implique donc que la variable ne participera plus à la variance et celle-ci diminuera d'autant plus que l'indice est élevé.

Les résultats présentés montrent un classement particulier pour chaque sortie. Cela confirme qu'il est possible d'intervenir de façon indépendante sur chaque entrée. D'autre part, les indices étant élevés seules les 3 premières variables auront un réel impact sur la modification du comportement de la sortie car responsables des principales variations.

Tableau III-8 : Classement des variables par influence pour chaque sortie prise indépendamment

Inconfort été	Parois opaques	Inconfort hiver	Besoins
Renouv. air (51)	Renouv. air (74)	Consigne (84)	Isolant Murs (57)
Isolant Vide S. (63)	Vitrage (84)	Isolant Murs (71)	Renouv. air (67)
Inertie (47)	Isolant Murs (92)	Vitrage (91)	Vitrage (79)
Vitrage (74)	Isolant Vide S. (86)	Isolant Vide S. (97)	Consigne (61)
Isol. Combles (50)	Inertie (96)	Isol. Combles (100)	Isolant Vide S. (92)
Isolant Murs (56)	Isol. Combles (64)	Renouv. air	Isol. Combles (92)
Ventil. non-hab. (44)	Consigne (100)	Ventil. non-hab.	Ventil. non-hab. (85)
Consigne	Ventil. non-hab.	Inertie	Inertie

La réalisation de ce type de tableau peut ainsi aider le concepteur dans sa vision du bâtiment en identifiant directement les principales variables qui déterminent la valeur d'une sortie et donc d'une performance. Observer les premières lignes du

tableau permet de visualiser rapidement le type d'opération qui aura un fort impact sur les différentes sorties observées.

3.3 Evaluation de la solution retenue

Une fois la solution retenue il est alors possible de réaliser une analyse d'incertitude pour visualiser le comportement local autour des entrées déterminées. Cette étape permet d'évaluer l'incertitude de sortie et donc de mener vers une possible garantie de performance.

Le Tableau III-9 suivant donne un récapitulatif des valeurs attribuées aux différentes variables et les lois et écarts associés. La gestion de ces information étant décrite plus en détail en Chapitre II :4.2. L'objectif est de déterminer l'influence des entrées sur la capacité à donner la bonne évaluation du résultat. Il est ainsi possible de déterminer sur quelles entrées un soin particulier devra être apporté.

Tableau III-9 : Solution retenue pour la réhabilitation et incertitudes associées

Elément	Unité	Moyenne	Incertitudes	Ecart
Technologie de Vitrages	Double PE		Discrète	–fixe–
Isolation des murs	[cm]	20	Gaussienne	±10%
Isolation du Vide Sanitaire	[cm]	20	Gaussienne	±10%
Isolation des combles	[cm]	20	Gaussienne	±10%
Renouvellement d'air / Infiltrations (ventilation)	[vol/h]	0,6	Gaussienne	±5%
Epaisseur des chapes (pour l'inertie)	[cm]	8	Gaussienne	±0,5cm
Ventilation des zones non-habitable Renouvellement d'air du vide sanitaire et des combles	[vol/h]	2	Gaussienne	±10%
Consigne de chauffage	[°C]	20,5	Uniforme	±0,5°C

Le Tableau III-10 montre les résultats obtenu, la même méthode que précédemment étant utilisée. Le renouvellement d'air intérieur reste toujours une variable très influente en été malgré le très faible écart défini, et cette fois l'isolation du vide sanitaire et l'inertie apparaissent également. En hiver, la consigne apparaît comme la principale variable, mais d'autres entrées apparaissent également telles que l'isolation des murs et le renouvellement d'air. Les effets sont ainsi plus mitigés mettant en avant d'autres variables influentes que dans les précédentes études. Les

principales variables, qui sont le renouvellement d'air et la température de consigne, sont généralement mal contrôlées durant la réhabilitation. Le Tableau III-10 les définit comme influentes, il est donc important de particulièrement bien définir ces entrées.

Tableau III-10 : Indices de sensibilité dans l'analyse d'incertitude

	Inconfort été	Parois opaques	Inconfort hiver	Besoins
Epaisseur d'isolant Murs	-0,18%	3,48%	-0,36%	11,04%
Ep. d'isolant Vide San.	10,56%	-0,38%	1,22%	-2,21%
Ep. d'isolant Combles	1,97%	1,67%	-1,89%	-3,07%
Inertie	8,40%	-0,60%	-0,04%	-1,47%
Renouvellement d'air	73,93%	93,28%	1,52%	9,03%
Ventilation z. non-hab.	-2,60%	0,62%	-0,44%	2,78%
Consigne de chauffage	2,46%	1,68%	99,82%	71,57%
Somme	94,53%	99,74%	99,82%	87,67%

Nous profitons de ces résultats pour présenter un comportement intéressant illustré en Figure III-19. L'amélioration du comportement d'hiver c'est-à-dire la réduction des besoins, mène à une dégradation des conditions estivales avec une élévation de la température intérieure. Ce problème est connu et généralement rencontré dans les simulations de bâtiments passifs, mais ne prend cependant pas en compte les actions des occupants qui auront alors tendance à agir sur les ouvrants pour améliorer leur confort. Le renouvellement d'air étant une variable essentielle, la mise en place d'un système de ventilation différent entre été et hiver permet de réduire cet effet. On comprend ainsi pourquoi des systèmes de sur-ventilation ou de ventilation traversante sont mis en place.

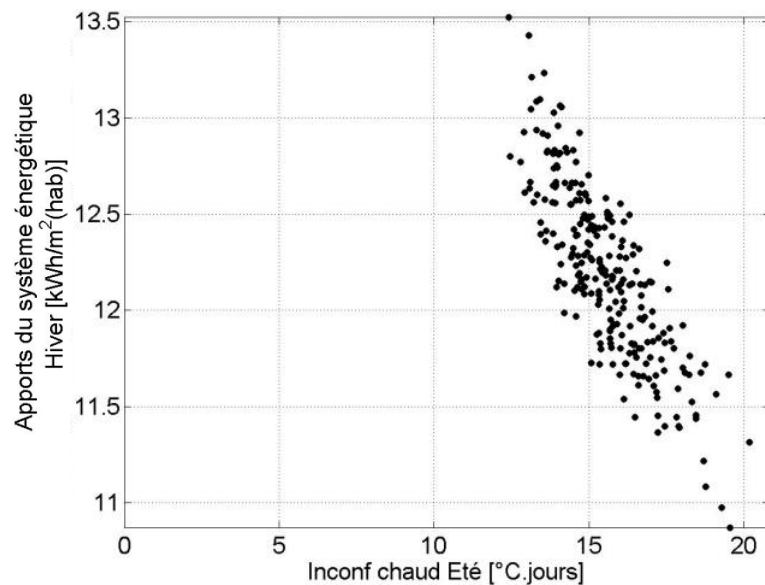


Figure III-19 : Evolution de l'inconfort en fonction de la réduction des besoins

3.4 Bilan de la méthodologie

La méthodologie présentée dans cette partie consiste donc à choisir des sorties d'intérêt et des variables qui pourront être modifiées par une opération. Le comportement du modèle est alors rapidement évalué par un criblage de l'espace des entrées. On connaît pour toutes les sorties le maximum de performance qu'il est possible d'atteindre en fonction des variables du modèle. Ce résultat n'est cependant pas directement utilisable puis qu'il correspond généralement à une solution non réalisable –exemple pour les besoins de chauffage : renouvellement d'air au minimum et isolation au maximum défini.

Une analyse de sensibilité à partir de ces résultats permet de mettre en avant et de quantifier le rôle des principales variables responsables des variations présentes en sortie. L'étude peut alors s'orienter vers l'analyse de ces comportements pour estimer rapidement les valeurs à attribuer aux entrées pour se rapprocher de la solution idéales.

Il est possible d'évaluer le comportement sur différentes saisons sachant que la problématique hiver est différente de celle d'été et que des comportements contradictoires peuvent donc apparaître.

La gestion des compromis n'est pas apparue dans cette étude, car les entrées ont des influences très différentes. Par exemple, la valeur de l'épaisseur d'isolant a été choisie uniquement en fonction des conditions d'hiver. Il aurait été possible de mettre en avant le fait que celle-ci participe à l'augmentation de l'inconfort en été, mais cet effet est négligeable face à l'incertitude de l'infiltration. Ce compromis disparaît ainsi, car

en cas de problème d'inconfort en été l'épaisseur d'isolant a un impact très faible par rapport aux autres entrées.

4 Une première étude encourageante mais des coûts de calcul trop important

Dans ce chapitre les objectifs du Chapitre I et les méthodes du Chapitre II ont été intégrés dans un programme. La thermique du bâtiment a besoin d'outils d'évaluation des incertitudes, de recherche de solution et d'analyse multicritères. L'outil présenté a pour objectif de pouvoir répondre à ces problématiques avec des méthodes issues de l'analyse de sensibilité. Dans une approche d'étude et d'optimisation globale du bâtiment, il semble préférable d'optimiser l'exploration de l'espace des réponses plutôt que d'optimiser une solution unique. Les simulations sont utilisées de façon différente et il est possible d'inclure des indicateurs économiques et environnementaux afin de pouvoir par la suite définir une solution.

En fonction des saisons, le comportement du bâtiment et les attentes sont différents avec parfois des comportements contraires. On est en présence de phénomènes complexes et l'optimisation automatique est donc particulièrement délicate. On retrouve ainsi une des conclusions du Chapitre I. L'analyse de sensibilité dans la recherche de solution de réhabilitation permet alors d'identifier rapidement les influences entrée-sortie. L'étude du système montre des relations d'amplitudes très différentes avec des évolutions non linéaires. Certains choix deviennent ainsi plus clairs et peuvent permettre de définir rapidement une estimation des performances atteintes quel que soit la valeur prise par les autres entrées. La variance de la sortie est réduite très rapidement en fixant les quelques entrées les plus influentes. Les choix sont réalisés élément par élément en fonction de leurs comportement et des valeurs atteintes sur les différentes sorties. Le choix de la solution est réalisé directement par le concepteur qui peut alors inclure naturellement tous les contraintes de projet qui sont généralement difficilement chiffrables.

Une fois une solution définie l'analyse de sensibilité joue également un second rôle en permettant d'analyser un comportement plus local du bâtiment. L'influence des sorties est alors différente et on remarque que certaines variables comme l'isolation permettent une augmentation des performances tout en ayant une influence relativement faible dans l'analyse d'incertitude en raison d'un comportement asymptotique. Il est alors possible d'affiner les choix et de donner un intervalle de confiance sur les performances atteintes vis-à-vis des variables considérées.

Différentes évaluations du même modèle avec les plages de variation différentes ont été réalisées. Le coût de calcul augmente alors rapidement bien que les évaluations soient toujours dans le même espace d'étude. D'autre part, l'analyse réalisée se fait obligatoirement en plusieurs étapes ce qui contraint l'utilisateur à interrompre sa réflexion pour réaliser de nouvelles simulations. L'objectif des travaux de thèse est alors de limiter le coût de calcul global que requiert l'analyse complète du bâtiment à réhabiliter.

Le nombre réduit de facteurs influents rend la recherche de solution manuelle réalisable à condition de disposer d'un affichage synthétique des différents comportements. Afin de répondre à la problématique du coût de calcul et de rendre accessible la recherche de solution manuelle, le Chapitre IV présente la construction d'un méta-modèle qui va permettre d'obtenir une estimation rapide des résultats du modèle dynamique sur un espace de variation spécifique. Celui-ci est complété par une interface associée qui permet de visualiser les informations sous forme synthétique.

Equation Chapter (Next) Section 1

Chapitre IV : CREATION D'UN META-MODELE POUR LA RECHERCHE DE SOLUTION

1 Objectifs du méta-modèle

Dans le chapitre précédent, nous avons vu l'outil développé pour la recherche de solution et l'analyse d'incertitude. Des applications ont été réalisées pour montrer son fonctionnement et comment il peut permettre d'analyser le comportement du modèle pour déterminer les valeurs à attribuer à chaque entrée. On remarque ainsi que seuls quelques facteurs de la réhabilitation thermique ont un impact important sur le comportement. L'optimisation peut être réalisée manuellement et permet un contrôle et une validation des choix réalisés. En contrepartie, l'étude peut nécessiter plusieurs étapes d'analyse. Dans un premier temps, on fait varier les entrées sur un large domaine d'étude pour rechercher les valeurs à attribuer aux principales entrées. Si une dispersion semble encore être présente, une étude supplémentaire est menée pour évaluer les variables de moins grande importance jusqu'à ce que les variations de chaque sortie soient acceptables. Puis dans un second temps, le comportement autour de la solution définie est étudié pour valider le résultat et déterminer quelles sont les variables qui risqueront d'invalidier le résultat lors de la mise en œuvre. La définition d'une solution doit coupler ces deux approches, il est alors nécessaire de réaliser de nombreux jeux de simulation pour la recherche et l'évaluation de l'incertitude correspondant à chaque solution retenue. L'analyse n'est alors plus immédiate et se décompose en de nombreuses phases de simulation puis recherche/contrôle. De plus, certaines zones sont évaluées à plusieurs reprises et il y a un nombre croissant de simulations, mais avec une valeur ajoutée relativement faible.

Il est alors nécessaire de retravailler la démarche pour rendre l'étude réalisable avec des temps de calcul raisonnables ou améliorer l'exploitation qui est faite de ses simulations. On peut pour cela :

- Réduire le modèle : une simulation plus simple est réalisée avec par exemple l'outil PHPP ou un modèle de parois de type RC.
- Construire un méta modèle : les résultats sont analysés pour reproduire un modèle réduit, plus léger, mais au comportement semblable sur un espace défini.

Les modèles simplifiés sont intéressants pour le coût de simulation mais ne peuvent réaliser que des évaluations de consommation moyenne et sont incapable de prendre correctement en compte les aspects solaires ou la régulation des systèmes. Ainsi, afin de rester dans la démarche du premier chapitre, le modèle complet dynamique est conservé et la solution du méta modèle est choisie. Ce type de reconstruction a déjà été réalisé pour le bâtiment, notamment à partir de relation linéaire ou des polynômes quadratiques issus d'un plan d'expérience (Jaffal et al., 2009). Il existe de nombreuses méthodes pour reconstruire un modèle :

- Régression linéaire : Très simple, mais forte hypothèse sur le comportement du modèle
- Représentation du modèle à haute dimension (HDMR – High Dimensional Model Representation) :
 - Approximation du modèle par une série composée des effets principaux et des interactions, mais tronquée à partir d'un certain ordre
 - Les coefficients de la suite sont approximés. Ex : méthode des splines, polynômes orthogonaux, fonction de base radiale.
- Représentation à tout ordre
 - Une surface de réponse est reconstruite avec les résultats des simulations. Ex : Krigeage, forêts aléatoires, Boosting.
 - Evolution exponentielle du coût de calcul avec le nombre de dimensions et les plages de variation des entrées.

Quel que soit la technique il est important de vérifier les hypothèses faites sur le comportement des sorties et la validité des nouveaux résultats. Garder un modèle de STD complet pour réaliser une recherche de solution sur de grandes plages de variations implique de choisir une méthode qui permette de reconstruire le comportement de sorties non linéaires sans demander un coût de calcul trop important. Ce type de reconstruction permet alors d'évaluer le bâtiment avec la même précision que le modèle de simulation thermique dont il est issu, mais sous

une forme qui peut être rapidement obtenue. La condition est de rester dans le domaine de définition pour lequel le méta modèle a été créé. Il est alors possible de réaliser une étude complète et rapide d'un modèle.

En analyse de sensibilité, la création d'une approximation du modèle peut ainsi être moins coûteuse en temps de calcul que la réalisation directe de l'étude des indices. Il existe un certain nombre de méthodes adapté à l'étude des indices, l'analyse se décompose alors en une phase d'apprentissage pour construire le modèle puis en une phase d'analyse. Les méthodes utilisées doivent donner des résultats de très bonne qualité afin de permettre une évaluation correcte des indices de sensibilité. Ce gain de performance est généralement permis, car le modèle initial doit être continu sur l'ensemble de l'espace d'étude. Les évaluations de celui-ci permettent de reconstruire la surface de réponse et d'optimiser l'analyse du comportement. Une fois le méta-modèle construit il est possible de réaliser des évaluations selon des variations statistiques particulières sur les entrées avec un coût de calcul négligeable par rapport au temps de simulation initial.

Comme dans la partie précédente, le modèle est considéré comme une boîte noire et il ne sera pas possible d'en extraire des informations autrement qu'avec des évaluations. La méthode du chaos polynomial répond alors parfaitement aux besoins de l'étude, car elle permet de produire un Meta modèle sans *a priori* sur la forme des variations et conserve donc la complexité du modèle initial. De plus elle permet une évaluation analytique de l'intégralité des indices de sensibilité, fournit des indicateurs d'erreur robustes et peut également produire un méta modèle qui dispose d'entrées discrètes.

La construction du méta modèle peut nécessiter plus d'appels au modèle qu'une simple analyse de sensibilité sur les effets principaux. Cependant l'objectif est d'ajouter une compréhension du modèle sur un domaine large et pas uniquement de déterminer une solution optimale. Les coûts de calcul restent tout de même acceptables, il est question de 1000 simulations soit 8h30 CPU pour des simulations de 30 secondes. Cependant, le temps de simulation peut fortement augmenter en fonction des modèles et aller jusqu'à 5 minutes si des modèles aérauliques sont présents et qu'une étude de l'éclairement intérieur est réalisée. Avec les puissances de calcul actuel, un ordinateur personnel multi-cœur peut parfaitement produire une analyse complète en une nuit ce qui laisse ensuite la possibilité à l'utilisateur de réaliser une étude sans coût de simulation supplémentaire.

Avec la possibilité de réaliser des évaluations avec un coût de calcul négligeable, le méta modèle permet d'étendre les possibilités de la recherche de solutions. En effet la configuration d'une méthode d'optimisation est trop complexe et influe de façon trop importante sur les résultats et il est proposé à la place de se servir du méta modèle pour avoir un retour rapide et fiable sur un ensemble de configuration avec pour chaque cas une évaluation de l'erreur, l'identification des variables responsable de cette erreur ainsi que le rôle général de chacune de ces variables dans les performances obtenues.

La partie suivante présente le principe de la décomposition en polynôme de chaos. Une version de cette méthode a été développée au sein du programme, mais a nécessité la mise en place d'algorithmes pour améliorer la qualité des résultats en ayant recours à une *toolbox* développé durant la thèse de G. Blatman (Blatman, 2009).

2 Le chaos polynomial

2.1 Principe de la décomposition en chaos polynomial – PCE

La méthode du chaos polynomial a été introduite dans les années 30, elle permet de projeter le signal de sortie dans une base de polynômes multi-variés dépendant de chaque entrée. Cette base de polynôme, généralement appelée PC pour Polynomial Chaos, est orthogonale ce qui permet de fournir une représentation unique du modèle dans cette base. La méthode originale aussi appelée chaos homogène est définie pour des variables gaussiennes et utilise uniquement des polynômes de Hermite comme base de l'espace aléatoire (Wiener, 1938). Cette définition a été généralisée à toutes les familles de polynômes orthogonaux et permet de démontrer que si une variable dispose d'une variance finie il est possible de représenter cette variable aléatoire dans une base de polynômes associée (Cameron et Martin, 1947; Xiu et Karniadakis, 2002). La décomposition en chaos polynomial d'un modèle quelconque \mathcal{M} est la suivante :

$$\mathcal{M}(x_1, \dots, x_m) = \sum_{i=0}^{\infty} a_i \psi_i(x_1, \dots, x_m) \quad (\text{IV.1})$$

Avec a_i des coefficients à estimer qui sont les coordonnées de la réponse du modèle dans la base polynomiale qui a été définie. La famille $\psi(x_1, \dots, x_n)$ est construite à partir de bases de polynômes uni-variés présélectionnées (Soize et Ghanem, 2004).

Pour chaque variable d'entrée, une base spécifique est choisie en fonction de la distribution de probabilité associée à la variable, elles peuvent être choisies à partir de l'ensemble des polynômes orthogonaux hypergéométriques d'Askey. Le Tableau

IV-1 présente quelques types de distribution avec les polynômes associés, leurs normes à l'ordre n et leurs intervalles de définition. Un choix correct de la famille de polynôme vis-à-vis de la distribution d'entrée permet de réduire la non-linéarité de la décomposition et améliore la convergence. Les bases des polynômes sont choisies afin que la fonction pondération des éléments orthogonaux corresponde à la densité de probabilité. Des bases discrètes sont également disponibles, avec les polynômes de Charlier qui sont associés à une distribution de poisson, mais non implémentées.

Tableau IV-1 : Distributions courantes et bases de polynômes orthogonaux associées

Distribution	Base polynomiale	Fonction de densité	Espace de Hilbert $\phi^{(j)}$	Support
Gaussienne	Hermite H_n	$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$	$H_n / \sqrt{n!}$	$] -\infty, +\infty[$
Uniforme	Legendre P_n	1	$P_n \sqrt{(2n+1)/2}$	$] -1, +1[$
Exponentielle	Laguerre $L_n^{\alpha=0}$	e^{-x}	L_n	$] 0, +\infty[$

La construction de la famille $\psi(x_1, \dots, x_m)$ se fait par le produit tensoriel de chaque base pour former toutes les combinaisons possibles entre les m bases. On introduit pour cela l'indice α_j^i qui détermine d'ordre du polynôme $\phi^{(j)}$ choisi dans la famille associée à la variable j .

$$\psi_i(x_1, \dots, x_m) = \prod_{j=1}^m \phi_{\alpha_j^i}^{(j)}(x_j) \quad (\text{IV.2})$$

Ainsi, on forme un Chaos Polynomial (PC) dont l'orthogonalité assure l'unicité de la décomposition et le calcul direct des indices de sensibilité grâce à la normalisation.

$$\langle \psi_i, \psi_j \rangle = \delta_{ij} \quad (\text{IV.3})$$

Cette suite de polynômes est tronquée pour des raisons de calcul numérique. La représentation suivante sera alors utilisée pour décrire le nouveau modèle construit à partir de la décomposition en chaos polynomial.

$$\mathcal{M}_{\text{PC}}(x_1, \dots, x_m) = \sum_{i=0}^{P-1} a_i \cdot \psi_i(x_1, \dots, x_m) \approx y \quad (\text{IV.4})$$

En considérant p le degré maximal des polynômes, on impose à l'équation (IV.2) la relation suivante :

$$\sum_{j=1}^m \alpha_j^i \leq p \quad \forall i \in \{0, 1, \dots, p-1\} \quad (\text{IV.5})$$

On autorise ainsi toutes les combinaisons entre polynômes de chaque base à condition que l'ordre final reste sous une limite. Le nombre d'éléments qui compose la suite est alors égal à :

$$P = \frac{(m+p)!}{m!p!} \quad (\text{IV.6})$$

Ne connaissant pas le modèle il est nécessaire de faire des hypothèses sur l'ordre maximum des polynômes et donc estimer sa complexité. Un polynôme d'ordre trop faible ne permet pas de représenter correctement la sortie. Mais choisir un ordre élevé implique de calculer un nombre très important de facteurs et si le nombre de simulations n'est pas suffisant on dégrade également la qualité de la réponse. Le choix de l'ordre maximal du PC produit donc un impact à la fois sur la capacité du modèle à représenter des fonctions complexes et sur le nombre d'inconnues à évaluer.

Pour illustrer l'effet de la complexification du modèle le Tableau IV-2 présente le nombre d'inconnues à calculer pour des valeurs croissantes de l'ordre p ou du nombre de variables m . Une exponentielle peut être correctement approximée par un polynôme d'ordre 5. Si on considère 6 variables d'entrées, le nombre de coefficients à calculer est de 462 et passe à 924 pour un ordre 6. Mais ce nombre augmente de façon importante avec la complexité.

Tableau IV-2 : Nombre d'inconnues à calculer en fonction de l'ordre du polynôme ou du nombre de variables

$m \backslash p$	1	2	3	4	5	6	7	8	9	10
1	2									
2	3	6								
3	4	10	20							
4	5	15	35	70						
5	6	21	56	126	252					
6	7	28	84	210	462	924				
7	8	36	120	330	792	1 716	3 432			
8	9	45	165	495	1 287	3 003	6 435	12 870		
9	10	55	220	715	2 002	5 005	11 440	24 310	48 620	
10	11	66	286	1 001	3 003	8 008	19 448	43 758	92 378	184 756

Des méthodes ont donc été mises en place pour réduire le nombre de coefficients. La plus simple est d'ajouter une contrainte supplémentaire sur les polynômes qui composent la base. Le degré des polynômes est limité en fonction de l'ordre des

interactions de façon à réduire de façon significative la taille de la base. C'est-à-dire qu'il est possible de donner le degré maximum aux polynômes uni-variés, et un degré de plus en plus réduit pour les polynômes à 2, 3, 4, ... variables.

Une alternative intéressante est offerte par la méthode LASSO (Least Absolute Shrinkage and Selection Operator), celle-ci permet d'ignorer certains coefficients afin de produire un chaos polynomial creux où les facteurs les moins influents sont exclus du calcul. Ainsi un modèle en x^3 ne disposera pas de polynômes de degré 1 ou 2 et le nombre de coefficients inconnus est significativement réduit sans altérer la qualité de la solution. Cette méthode permet alors d'utiliser une décomposition en chaos polynomial de plus haut degré sans forcément augmenter le nombre d'inconnues.

2.2 Exemple de décomposition simple et lien avec les indices de sensibilité

Prenons par exemple le cas de deux variables aléatoires définies respectivement par une loi gaussienne et uniforme. Les bases polynômiales de Hermite et Legendre seront utilisées pour correspondre aux lois de distribution des entrées, comme présenté dans le Tableau IV-1. Les polynômes de Legendre sont définis sur un espace réduit, il est donc nécessaire de réaliser un changement de variable afin d'adapter le domaine de variation des entrées au support des polynômes. A la suite de ce changement de variable, on définit le domaine de variation D_x qui correspond à l'ensemble des points qui seront explorés par la variable $\mathbf{x} = \{x_1, x_2, \dots, x_m\}$.

L'équation suivante présente les premiers termes des bases polynomiales de Hermite et Legendre qui serviront pour la construction du chaos polynomial :

$$\begin{array}{l|l} H_0 = 1 & P_0 = 1 \\ H_1 = x & P_1 = x \\ H_2 = (x^2 - 1)/\sqrt{2} & P_2 = (3x^2 - 1)\sqrt{5/6} \end{array} \quad (IV.7)$$

Le PC construit jusqu'à l'ordre 2 avec l'équation (IV.2) est alors composé de 6 polynômes :

$$\begin{array}{l|l} \psi_0 = H_0^{(1)} \cdot P_0^{(2)} = 1 & \psi_3 = H_2^{(1)} \cdot P_0^{(2)} = (x_1^2 - 1)/\sqrt{2} \\ \psi_1 = H_1^{(1)} \cdot P_0^{(2)} = x_1 & \psi_4 = H_1^{(1)} \cdot P_1^{(2)} = x_1 \cdot x_2 \\ \psi_2 = H_0^{(1)} \cdot P_1^{(2)} = x_2 & \psi_5 = H_0^{(1)} \cdot P_2^{(2)} = (3x_2^2 - 1)\sqrt{5/6} \end{array} \quad (IV.8)$$

Ainsi le modèle polynomial complet pour cette base s'écrit :

$$y \approx a_0 + a_1 x_1 + a_2 x_2 + a_3 (x_1^2 - 1)/\sqrt{2} + a_4 x_1 x_2 + a_5 (3x_2^2 - 1)\sqrt{5/6} \quad (IV.9)$$

Comme pour les méthodes ANOVA (Chapitre II :2.1) il est possible de faire apparaître explicitement les contributions sur la sortie de chaque variable d'entrée et de toutes leurs combinaisons. Pour rappel, dans cette décomposition le modèle est représenté par : f_0 la moyenne, $f_i(x_i)$ les contributions des variables seules et $f_{i...k}(x_i, \dots, x_k)$ les contributions des variables par interaction.

$$\mathcal{M}(\mathbf{x}) = f_0 + \sum_{i=1}^m f_i(x_i) + \sum_{i < j} f_{ij}(x_i, x_j) + \dots + f_{1...m}(x_1, \dots, x_m) \quad (\text{IV.10})$$

Les polynômes sont ainsi regroupés de la façon suivante :

$$\begin{aligned} f_0 &= a_0 \\ f_j(x_j) &= \sum_{i \in \alpha_j} a_i \cdot \psi_i(x_1, \dots, x_m) \\ f_{jk}(x_j, x_k) &= \sum_{i \in \alpha_{j,k}} a_i \cdot \psi_i(x_1, \dots, x_m) \end{aligned} \quad (\text{IV.11})$$

Tel que α_j le multi-indice ciblant tous les polynômes du PC faisant intervenir uniquement la variable j , de même $\alpha_{j,k}$ le multi-indice faisant intervenir uniquement les interactions entre la variable j et la variable k et ainsi de suite jusqu'à l'interaction maximale. Appliqué à l'exemple actuel, on identifie donc les contributions suivantes :

$$\begin{aligned} f_0 &= a_0 \\ f_1(x_1) &= a_1 x_1 + a_3 (x_1^2 - 1) / \sqrt{2} \\ f_2(x_2) &= a_2 x_2 + a_5 (3x_2^2 - 1) \sqrt{5/6} \\ f_{12}(x_1, x_2) &= a_4 x_1 x_2 \end{aligned} \quad (\text{IV.12})$$

Dans le Chapitre II :2.1 à l'équation (II.7) nous avons vu que cette décomposition permettait le calcul des indices de sensibilité en calculant le rapport entre la variance de chaque fonction sur la variance totale du modèle. Appliqué au modèle polynomial cela revient à évaluer la variance de chacun des polynômes qui compose le modèle. Par exemple pour le calcul de la variance totale on a :

$$\begin{aligned} V &= \int_{D_x} (\mathcal{M}_{PC}(\mathbf{x}) - \bar{y})^2 p_x(\mathbf{x}) d\mathbf{x} \\ &= \int_{D_x} \left(\sum_{i=1}^{p-1} a_i \cdot \psi_i(\mathbf{x}) \right)^2 p_x(\mathbf{x}) d\mathbf{x} \end{aligned} \quad (\text{IV.13})$$

Avec $p_x(\mathbf{x})$ une fonction de poids généralement appelée fonction de densité de probabilité jointe qui permet de représenter la répartition aléatoire des entrées lors des essais, une définition plus détaillée est présente à l'équation (II.2).

Cependant la forme des polynômes est connue, ils sont normés et sélectionnés pour correspondre à la distribution des variables d'entrée. Le produit scalaire de deux polynômes dans l'espace des variables est donc égal à zéro ou un :

$$\langle \psi_i(\mathbf{x}), \psi_j(\mathbf{x}) \rangle = \int_{D_x} \psi_i(\mathbf{x}) \cdot \psi_j(\mathbf{x}) \cdot p_x(\mathbf{x}) d\mathbf{x} = \delta_{ij} \quad (\text{IV.14})$$

La combinaison des équations (IV.13) et (IV.14) permet alors d'obtenir une relation directe entre la variance et les coefficients des polynômes.

$$V = \sum_{i=1}^{P-1} a_i^2 \quad (\text{IV.15})$$

Cette relation peut être étendue aux différentes variances conditionnelles pour former les indices de sensibilité. On obtient ainsi une évaluation analytique, et donc sans erreur, des indices à tout ordre :

$$S_j = \frac{\sum_{i \in \alpha_j} a_i^2}{V} \quad (\text{IV.16})$$

$$S_{1,\dots,j} = \frac{\sum_{i \in \alpha_{1,\dots,j}} a_i^2}{V}$$

Ces indices sont cependant ceux du modèle polynomial et non pas ceux du modèle initial. Les coefficients subissent une approximation due à la suite tronquée et également lors de leur identification.

Il est donc important de réaliser un calcul correct des coefficients et ensuite de fournir une évaluation de l'erreur résiduelle.

2.3 Calcul des coefficients des polynômes

Le calcul des coefficients peut se faire de différentes manières. En ne considérant que les méthodes non-intrusives, les deux approches principales sont par projection et régression (Blatman et Sudret, 2010).

La nature orthogonale du PC permet d'obtenir les coefficients en réalisant une projection sur chacun des éléments de la base. Le calcul de ces coefficients requiert alors d'évaluer le produit scalaire de la sortie avec un polynôme associé :

$$\langle \mathcal{M}_{PC}(\mathbf{x}), \psi_i(\mathbf{x}) \rangle = a_i \quad (\text{IV.17})$$

Le produit scalaire est défini à l'équation (IV.14) et nécessite la résolution d'une intégrale qui ne dispose pas de forme analytique. Celle-ci doit alors être évaluée par des méthodes de quadrature qui ne seront pas détaillées ici. Ces méthodes sont très performantes, mais le coût de calcul devient très important avec l'augmentation du nombre de coefficients à calculer.

L'autre approche possible est de chercher les coefficients par régression en minimisant l'écart entre le modèle polynomial et le modèle étudié. On écrit pour cela la décomposition sous forme matricielle avec :

$$\begin{aligned} \mathbf{a} &= \{a_0, \dots, a_{p-1}\}^T \\ \boldsymbol{\psi}(\mathbf{x}) &= \{\psi_0(\mathbf{x}), \dots, \psi_{p-1}(\mathbf{x})\} \end{aligned} \quad (\text{IV.18})$$

Et on pose le problème en évaluant l'erreur quadratique sur un nombre N de réalisations :

$$\mathbf{a} = \arg \min \sum_{n=1}^N \left(\mathbf{a} \cdot \boldsymbol{\psi}(\mathbf{x}^{(n)}) - \mathcal{M}(\mathbf{x}^{(n)}) \right)^2 \quad (\text{IV.19})$$

Cela revient à poser l'équation (IV.20) avec $\boldsymbol{\Psi}$ une matrice de dimension N par P , qui correspond aux P polynômes de la base exprimés pour chacune des N combinaisons de valeurs d'entrées. Et \mathbf{Y} , un vecteur colonne qui contient les résultats des N réalisations.

$$\mathbf{a} \cdot \boldsymbol{\Psi} = \mathbf{Y} \quad \begin{cases} \boldsymbol{\Psi} = \{\boldsymbol{\psi}(\mathbf{x}^{(1)}), \dots, \boldsymbol{\psi}(\mathbf{x}^{(N)})\}^T \\ \mathbf{Y} = \{\mathcal{M}(\mathbf{x}^{(1)}), \dots, \mathcal{M}(\mathbf{x}^{(N)})\}^T \end{cases} \quad (\text{IV.20})$$

Il est ainsi nécessaire d'inverser $\boldsymbol{\Psi}$ qui peut se mettre sous la forme d'une matrice symétrique avec la transformation suivante :

$$\mathbf{a} = \mathbf{Y} \cdot \boldsymbol{\Psi}^T (\boldsymbol{\Psi} \cdot \boldsymbol{\Psi}^T)^{-1} \quad (\text{IV.21})$$

La qualité de la solution dépend alors du conditionnement de $\boldsymbol{\Psi} \cdot \boldsymbol{\Psi}^T$ et par conséquent du nombre de simulations. Le nombre de simulations nécessaires doit

normalement être au moins égal au nombre d'inconnues afin d'être sûr de disposer d'un problème bien posé mais il est possible de réduire le nombre de simulations. Le problème devient alors mal posé et n'assure plus l'unicité de la solution, car il peut exister plusieurs jeux de valeur qui permettent au modèle polynomial de passer par les points évalués et le comportement entre les points n'est ainsi plus maîtrisé et peut produire des valeurs aberrantes. Ce comportement peut être pallié par une méthode de régularisation, donnée à l'équation (IV.22), qui ajoute une contrainte sur les coefficients et peut ainsi permettre de retrouver un problème bien posé.

$$\sum |a_i| \leq s \quad (\text{IV.22})$$

Des méthodes de régression plus évoluées peuvent alors être employées. Notamment la méthode de résolution LAR (Least Angle Regression) qui permet une recherche des coefficients par ordre d'influence (Efron et al., 2004). L'initialisation place tous les coefficients à zéros et définit le résidu de l'analyse comme les variations de la sortie. Ce résidu est alors comparé avec la base polynômiale pour déterminer avec quel polynôme $\psi_i(\mathbf{x})$ le résidu est le plus corrélé. Le coefficient de ce polynôme passe alors de zéro à la valeur donnée par le coefficient des moindres carrés entre ce polynôme et le résidu actuel, jusqu'à ce qu'un autre polynôme ait une plus forte corrélation. Les coefficients de ces deux polynômes sont alors déplacés ensemble dans la direction donnée par le coefficient des moindres carrés calculé avec les deux polynômes et le nouveau résidu, jusqu'à ce qu'un troisième coefficient soit introduit. L'analyse se poursuit alors jusqu'à ce que tous les coefficients aient été considérés ou que le nombre de coefficient soit égal au nombre de simulations.

Associé à la méthode LASSO précédemment décrite il est possible de ne choisir que les coefficients utiles à la représentation du modèle initial. La mise en place d'un algorithme adaptatif, généralement appelé LARS (où S peut correspondre à Stagewise ou à LASSO) permet de faire évoluer le choix des coefficients pour tendre vers la meilleure solution. Cette recherche de solution nécessite alors un estimateur d'erreur pour évaluer ces choix. Celui-ci est décrit dans la section suivante.

Dans tous les cas, la bonne convergence de ces méthodes ne peut se faire que si une bonne exploration de l'espace est réalisée. Il est alors fortement recommandé d'utiliser des méthodes d'échantillonnage à faible discrétance type LP-Tau ce qui a été détaillé dans le Chapitre II :4.3.2.

2.4 Indicateur d'erreur

L'évaluation de l'erreur étant importante, il est nécessaire d'utiliser un estimateur d'erreur performant pour permettre de construire et de valider le modèle polynomial. Classiquement, une approximation de l'erreur est produite par le calcul via la norme L^2 de la distance moyenne entre le modèle et son approximation. Cette erreur est alors divisée par la variance des résultats pour fournir une évaluation adimensionnelle. Appliquée à un échantillon de N réalisations on obtient :

$$E_{exp} = \frac{1}{N} \sum_{n=1}^N \left(\mathcal{M}_{PC}(x^{(n)}) - \mathcal{M}(x^{(n)}) \right)^2$$

$$\varepsilon_{exp} = \frac{E_{exp}}{V}$$
(IV.23)

Une notation généralement utilisée appelée R^2 consiste à représenter le complément à 1 de l'erreur ε_{exp} :

$$R^2 = 1 - \varepsilon_{exp}$$
(IV.24)

Cependant, une estimation correcte de l'erreur doit faire appel à une validation croisée, c'est-à-dire qu'il est nécessaire d'évaluer l'erreur sur des points indépendants qui n'ont pas été utilisés. Si l'erreur est calculée avec les points qui ont servi à construire le modèle, l'évaluation est faussée et on produit un estimateur biaisé. Le principal problème est que la recherche des coefficients demande déjà un nombre important de simulations et il est difficile de disposer de simulations supplémentaires pour évaluer l'erreur commise en différents points sans que celles-ci soient utilisées.

Il existe alors un type de méthode connue sous le nom de *Leave One Out* qui permet d'obtenir une évaluation correcte sans simulations supplémentaires. Cette erreur est évaluée à partir de l'écart entre un essai et le méta modèle construit à partir des $N-1$ autres essais restant. En calculant la moyenne de l'erreur sur chacun des points, il est possible de simuler une évaluation croisée sans simulations supplémentaires. On obtient ainsi une évaluation performante de l'erreur sur l'ensemble du domaine appelée PRESS (Predicted Residual Sum of Squares) :

$$E_{PRESS} = \frac{1}{N} \sum_{n=1}^N \left(\mathcal{M}_{PC \setminus n}(x^{(n)}) - \mathcal{M}(x^{(n)}) \right)^2$$
(IV.25)

Avec $\mathcal{M}_{PC \setminus n}$ le modèle polynomial construit sans la n -ième évaluation.

Comme précédemment une transformation est réalisée pour obtenir une définition plus classique appelée Q^2 .

$$Q^2 = 1 - \frac{E_{PRESS}}{V} \quad (IV.26)$$

Cette évaluation permet alors d'évaluer la robustesse de la méthode, mais également la capacité du modèle à évaluer les autres points de son domaine d'étude.

Un avantage non négligeable de cet estimateur est qu'il peut s'obtenir directement à partir du modèle polynomial complet et de la matrice $(\Psi \cdot \Psi^T)^{-1}$ qui a servi à sa construction. L'évaluation de cette erreur ne demande donc pas le calcul d'autres régressions et peut être employée dans un algorithme pour la recherche des coefficients comme LARS (Blatman et Sudret, 2009) présenté précédemment.

3 Comparaisons avec le modèle réel

Une évaluation du méta-modèle sur un modèle de simulation thermique dynamique va maintenant être réalisée. Pour cela le modèle de la maison INCAS, présenté dans le Chapitre III :3.1, est utilisé comme base. Les quatre sorties précédemment décrites : l'inconfort en été, l'énergie entrante des parois opaques en été, l'inconfort en hiver et les besoins en hiver ; sont évaluées à tour de rôle pour former quatre modèles polynomiaux indépendants. Dans la suite ces quatre modèles seront confondus sous l'appellation modèle polynomial par opposition au modèle dynamique.

La démarche est en tout point identique à une analyse de sensibilité à la différence qu'un échantillonnage de type LP-Tau –voir Chapitre II :4.3.2– est utilisé pour avoir une meilleure convergence et 1000 simulations sont réalisées afin de fournir suffisamment de points d'évaluation. Le modèle polynomial est alors calculé à partir de l'algorithme adaptatif basé sur la méthode LAR précédemment décrite.

D'autre part, un second jeu de 1000 simulations est réalisé avec un échantillonnage LHS. Ces résultats serviront de référence afin d'évaluer la qualité du modèle polynomial construit sur des points indépendants.

Cette comparaison s'appuiera sur différentes méthodes d'évaluation numérique ou qualitative, afin de fournir une estimation des capacités de ce type de représentation.

3.1 Calcul d'erreur et convergence

Dans un premier temps trois évaluations d'erreur sont calculées.

- ε_{PRESS} l'erreur issue de la construction du méta modèle ;
- ε_{locale} l'erreur locale qui est mesurée avec les points qui ont servi à la construction du méta modèle ;
- $\varepsilon_{croisée}$ l'erreur croisée qui est mesurée à partir du second jeu de 1000 essais (LHS).

Les informations relatives à ces indicateurs sont présentes en section 2.4. D'après leurs définitions ceux-ci doivent tendre vers zéro, cette valeur est cependant une limite qui ne peut être atteinte si le modèle présente certaines discontinuités.

Les Figure IV-1 et Figure IV-2 rassemblent les évolutions des erreurs pour les quatre sorties en fonction du nombre de simulations réalisées. On remarque que les valeurs atteintes par l'algorithme pour les 1000 simulations sont très satisfaisantes, celles-ci varient entre 10^{-2} et 10^{-4} soit des coefficients de corrélation Q^2 entre 0,99 et 0,9999. L'estimation réalisée par cet indicateur est de meilleure qualité et plus stable que l'erreur locale qui produit effectivement une sur-estimation. La valeur reste cependant assez éloignée de l'erreur croisée qui est considérée comme l'évaluation de référence.

Cette erreur croisée converge d'ailleurs très rapidement alors que le meta-modèle semble nécessiter plus de simulation pour atteindre une erreur stable. On remarque d'ailleurs l'absence de phénomène de sur-aprentissage car l'erreur croisée reste stable alors qu'il y a une complexification du modèle polynomial. Pour rappel le phénomène de sur-aprentissage correspond à une dégradation générale de la capacité de prédiction du modèle malgré une amélioration sur les points qui servent à la construction de celui-ci. L'évaluation de la prédiction faite sur des points indépendants permet alors de détecter ce phénomène.

La convergence se fait donc avant les 200 simulations sauf pour l'inconfort en été qui est atteint vers 400. En fonction des sorties considérées, un nombre très réduit de simulations peut donc suffir, mais cette convergence va dépendre de la forme de la sortie, des entrées présentes et des interactions du modèle. Il est alors difficile d'estimer *a priori* le nombre de simulations nécessaire. Actuellement, un nombre arbitraire de simulations est choisi, mais à terme il est possible de réaliser un enrichissement progressif en fonction de l'erreur atteinte.

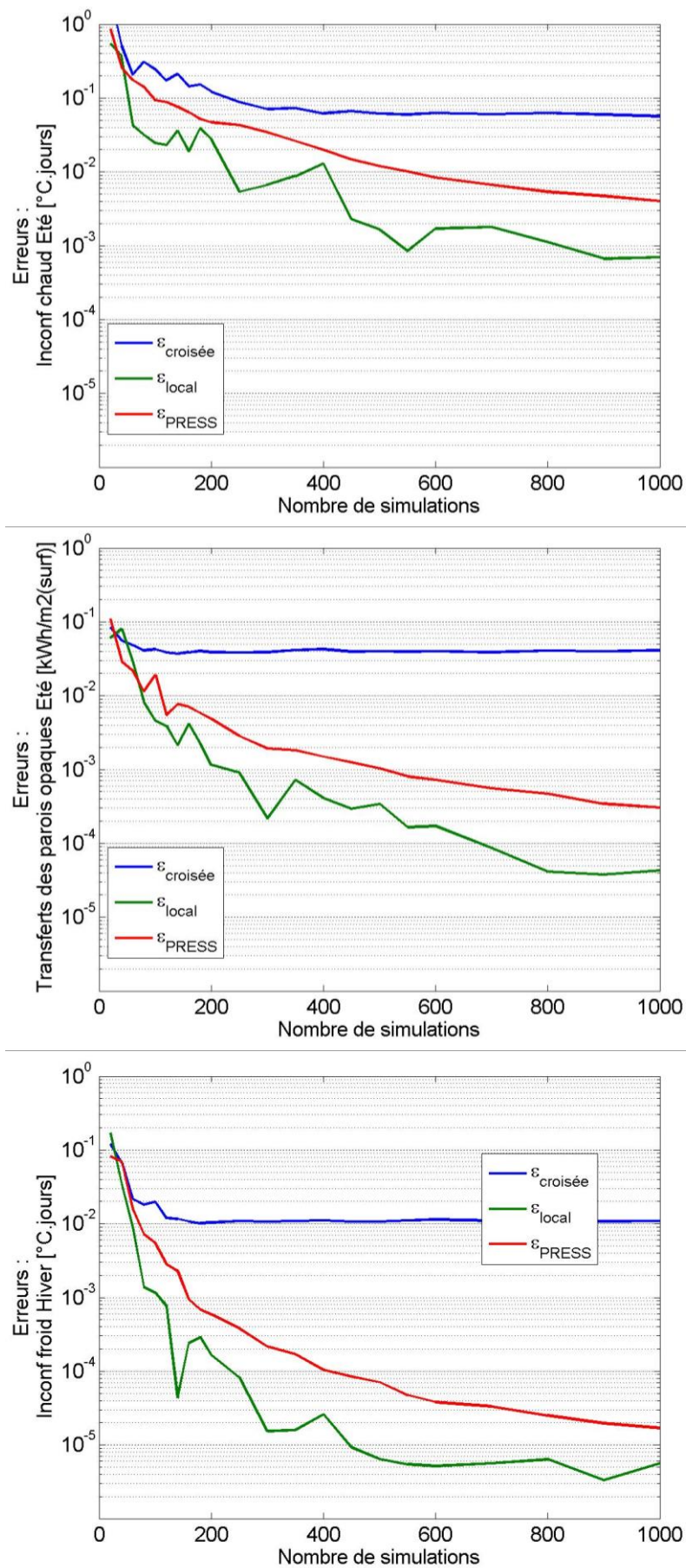


Figure IV-1 : Convergence des erreurs du méta-modèle -3/4

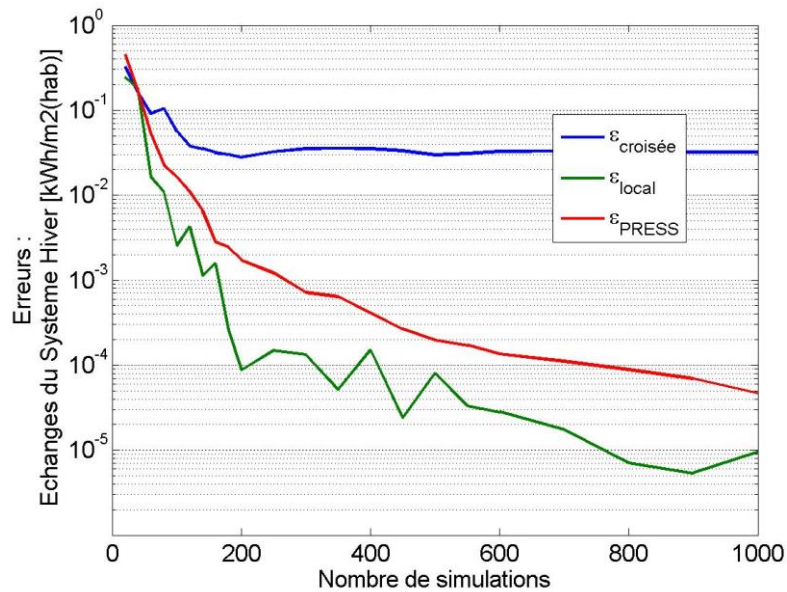


Figure IV-2 : Convergence des erreurs du méta-modèle -4/4

3.2 Comparaisons qualitatives

Une autre façon plus visuelle d'évaluer la qualité du méta-modèle obtenu est de tracer le *Scatterplot* des évaluations du modèle polynomial en fonction des résultats du modèle dynamique. Cette représentation est qualitative, elle permet de localiser les erreurs les plus probables et de visualiser l'écart par rapport à la diagonale.

La Figure IV-3 présente ci-dessous affiche les 1000 évaluations de référence en fonction de la valeur correspondante estimée par le méta-modèle. Le modèle polynomial utilisé est celui issu du calcul avec les 1000 essais. La Figure IV-4 présente le même résultat, mais issu du modèle réalisé avec seulement 300 simulations. Les différences entre les graphiques sont minimales et dans l'ensemble les essais sont relativement bien alignés mis à part une dispersion plus importante pour l'inconfort d'été conformément aux erreurs évaluées précédemment. Ce graphique vérifie ainsi la bonne corrélation entre les deux modèles et permet de constater que certains points sont éloignés de la diagonale, mais ne représente que des cas particuliers sur les 1000 simulations. La corrélation entre ces deux modèles est l'exacte définition de l'erreur croisée exprimée à l'équation (IV.23).

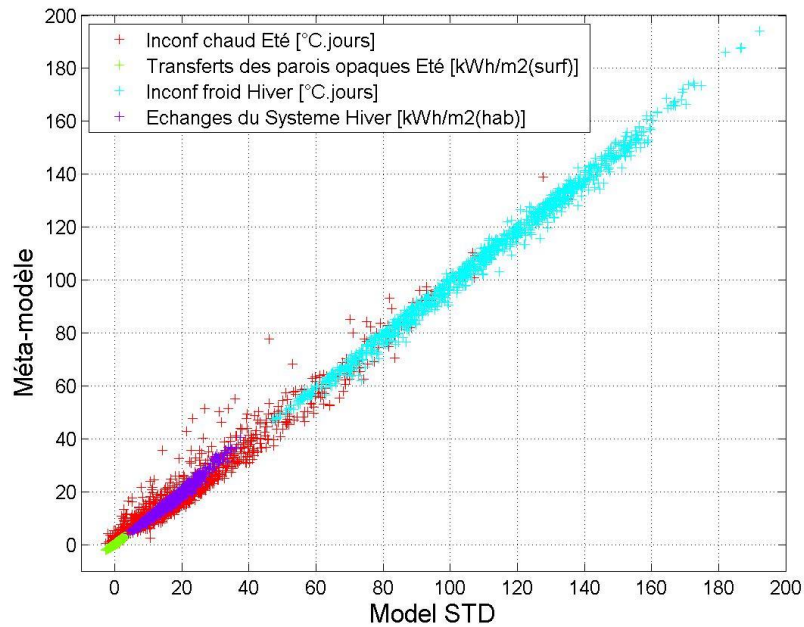


Figure IV-3 : Corrélation entre modèle STD et modèle polynômial (1000 simulations)

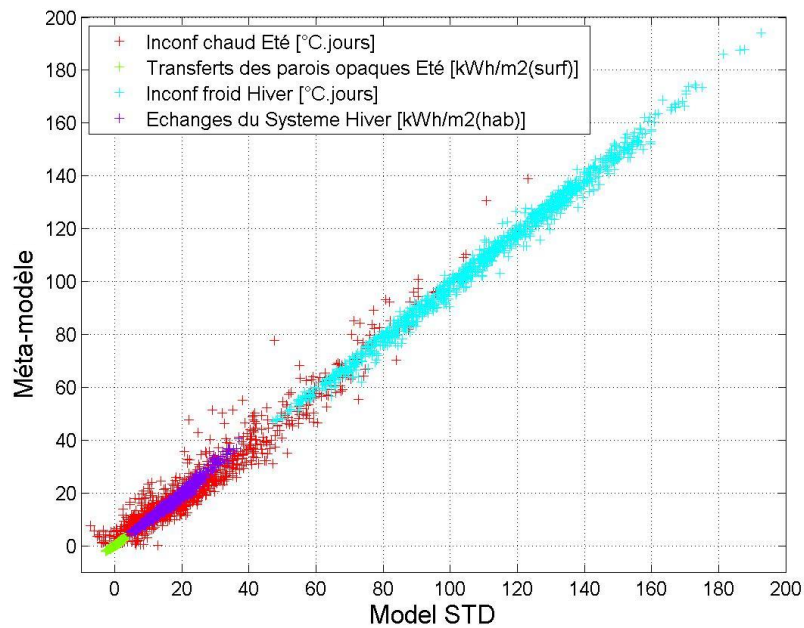


Figure IV-4 : Corrélation entre modèle STD et modèle polynômial (300 simulations)

L'observation des distributions de sortie, présentes en Figure IV-5, peut être un autre moyen d'évaluer ces différences entre modèle et méta-modèle. Ce type de représentation permet de réaliser une estimation efficace du risque de sous ou sur-évaluation des différentes sorties. Elle permet, d'autre part, de vérifier la bonne estimation des moyennes et variances.

Des formes complexes sont générées par le modèle et l'approximation polynomiale conduit à de légers décalages. Néanmoins, on peut s'attendre à des estimations

correctes et à chaque fois une légère surestimation –en moyenne– du modèle polynomial qui ainsi ne produit pas de risque de défaillance.

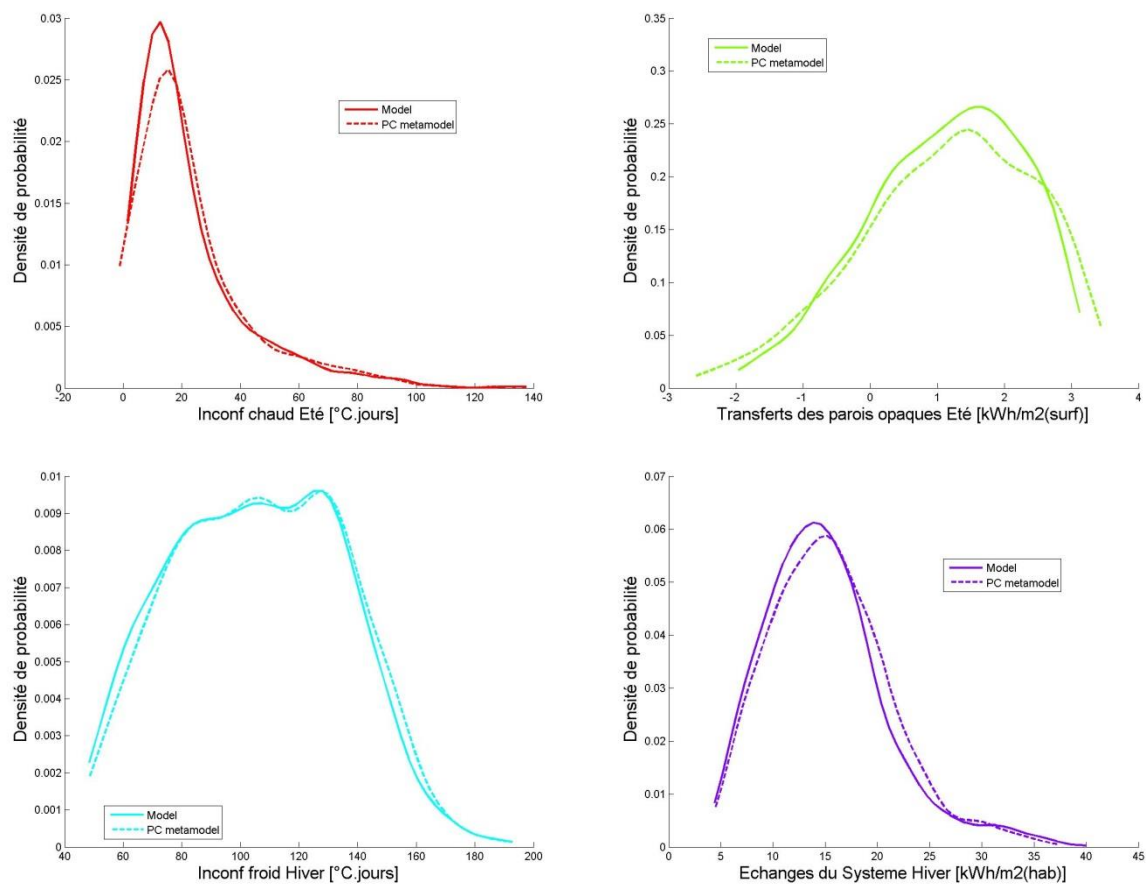


Figure IV-5 : Comparaison des distributions entre modèle et méta-modèle

Le programme permet ainsi d'obtenir des informations correctes pour l'analyse d'incertitude et l'analyse de sensibilité.

3.3 Indices de sensibilité

L'analyse de sensibilité pose en théorie quelques problèmes à cause de la présence d'entrées discrètes. Etant donné qu'aucun polynôme discret n'a été implémenté, la variable discrète est représentée par un polynôme pour une loi uniforme qui n'est évaluée qu'à des points particuliers. Le méta-modèle est ainsi correct si aucune évaluation n'est réalisée en dehors de ces points. Cependant lors de la construction les coefficients ne sont pas contraints en dehors des points d'évaluation, ils peuvent ainsi prendre des valeurs incorrectes. L'indice de sensibilité ainsi calculé peut être faussé vis-à-vis de l'effet réel de la variable. La méthode LAR assure cependant une certaine stabilité, car la valeur attribuée à chaque coefficient des polynômes augmente progressivement, les empêchant donc de prendre des valeurs aberrantes.

Le Tableau IV-3 et le Tableau IV-4 présentent respectivement les indices principaux et totaux calculés pour chaque sortie via les coefficients des polynômes comme décrits en section 2.2.

Le principe général de ces indicateurs est décrit en Chapitre II :2.1, l'indice principal d'une variable est l'effet de la variable seule et l'indice total correspond à l'effet de la variable avec ses interactions avec les autres variables. La somme de ces différents effets comme présentée à l'équation (II.13), permet une analyse rapide de l'importance des interactions présentes

Le Tableau IV-3 présente les indices principaux et montre des résultats très proches de ceux trouvés lors de l'analyse précédemment réalisée avec ce modèle et dont les résultats ont été présentés dans le Tableau III-6 en page 101. On confirme ainsi la précédente observation avec un échantillonnage, des simulations et une méthode d'analyse différente. La technique du chaos polynomial est cependant de plus haut niveau et permet une précision beaucoup plus importante dans l'évaluation des indices.

Tableau IV-3 : Indices de sensibilité principaux par chaos polynomial

	Inconfort été	Parois opaques	Inconfort hiver	Besoins
Epaisseur d'isolant Murs	0,88%	0,94%	11,8%	40,9%
Ep. d'isolant du Vide San.	27,2%	0,97%	1,56%	3,50%
Ep. d'isolant des Combles	0,01%	0%	0%	0%
Epaisseur des dalles (Inertie)	2,30%	0%	0%	0%
Renouvellement d'air	38,1%	62,9%	0%	29,9%
Ventilation non-habitable	0,01%	0%	0,01%	0,02%
Technologie de vitrage	7,41%	31,97%	6,52%	21,9%
Consigne de chauffage	0%	0%	80,01%	3,47%
Somme (indices principaux)	76,0%	96,8%	99,9%	99,6%

La somme des indices principaux est ici très proche de 100% sauf pour l'inconfort d'été. Cette remarque a déjà été faite précédemment sans pouvoir être vérifiée. Cependant la technique employée ici permet le calcul immédiat des effets à tout ordre et permet donc de produire les indices totaux présentés dans le Tableau IV-4. Celui-ci confirme que l'inconfort d'été est bien sujet à des interactions, car la somme est ici de 127%.

Tableau IV-4 : Indices de sensibilité totaux par chaos polynomial

	Inconfort été	Parois opaques	Inconfort hiver	Besoins
Epaisseur d'isolant Murs	6,55%	1,29%	11,9%	41,0%
Ep. d'isolant du Vide San.	41,4%	3,77%	1,61%	3,60%
Ep. d'isolant des Combles	0,04%	0,01%	0%	0%
Epaisseur des dalles (Inertie)	4,69%	0,03%	0%	0%
Renouvellement d'air	58,5%	64,8%	0,01%	29,9%
Ventilation non-habitable	0,05%	0,01%	0,02%	0,04%
Technologie de vitrage	16,2%	33,2%	6,60%	22,08%
Consigne de chauffage	0,02%	0,02%	80,04%	3,77%
Somme (indices totaux)	127%	103%	100%	100%

Pour connaître quelles sont ces interactions, il est nécessaire de détailler une partie des effets, le Tableau IV-5 présente les indices d'ordre deux pour l'inconfort d'été. Ces valeurs correspondent aux indices de sensibilité dus aux effets conjoints des variables prises deux à deux. On observe ainsi que le renouvellement d'air intérieur est le plus sujet aux interactions d'ordre deux et produit des effets avec les vitrages, les épaisseurs d'isolation des murs et du vide sanitaire. La somme de tous les effets du tableau donne 20,7%, on peut donc estimer que les interactions d'ordres supérieurs sont responsables des 3,3% d'effet restant.

Tableau IV-5 : Indices de sensibilité d'ordre deux par chaos polynomial pour l'inconfort d'été

Inconfort d'été		1	2	3	4	5	6	7
1	Epaisseur d'isolant Murs							
2	Ep. d'isolant du Vide San.	0,74%						
3	Ep. d'isolant des Combles	0%	0%					
4	Inertie	0,17%	0,55%	0%				
5	Renouvellement d'air	2,81%	8,99%	0%	0,87%			
6	Ventilation non-habitable	0%	0%	0%	0%	0%		
7	Technologie de vitrage	0,32%	1,48%	0%	0,21%	4,54%	0,01%	
8	Consigne de chauffage	0%	0%	0%	0%	0%	0%	0,01%

Cette étude est également réalisable avec les techniques présentées au Chapitre II :2.2 comme la méthode Sobol, mais nécessite un jeu de simulation pour chacune des 28 cases du Tableau IV-5. Elle peut être tout de même réduite à quatre variables en comparant les écarts entre le tableau indice principal et le tableau indice total. Ces quatre variables forment 6 combinaisons plus le calcul des indices totaux, en

considérant 500 simulations par étude il y a donc 3500 simulations supplémentaires à réaliser.

Le chaos polynomial est ainsi une méthode très performante pour évaluer les effets sans que le nombre de simulations soit trop important. Il permet au besoin de réaliser une étude poussée qui amène plus de compréhension sur les phénomènes en interaction. Cependant, réaliser une étude de sensibilité au sein du domaine de définition initial nécessite la réalisation d'un nouveau modèle polynomial sur le domaine réduit. Une nouvelle évaluation des coefficients est alors nécessaire et cette d'optimisation peut en fonction de la complexité durer de 7 secondes à 6 minutes.

Ce mode d'évaluation des indices ne sera donc pas retenu et une évaluation plus sommaire, mais très rapide sera préférée dans la recherche de solution. L'analyse de sensibilité complète avec : l'échantillonnage, l'évaluation par le méta-modèle et analyse ; prend un temps inférieur à 0,2sec avec RBD-FAST ou RBD-SOBOL. Ce type de méthode répond alors correctement aux besoins d'une interface utilisateur réactive.

4 Utilisation du méta-modèle – Interface utilisateur

Les résultats fournis sont une approximation du modèle initial, mais sont tout de même de bonne qualité. Le méta modèle offre une méthode efficace pour étudier le comportement à moindre coût, même si quelques valeurs peuvent donner des résultats incorrects.

Le modèle polynomial n'est pas destiné à remplacer le modèle dynamique, mais plutôt à simplifier l'approche de simulation pour aider l'utilisateur à analyser le comportement des entrées vis-à-vis des sorties. Il est à noter que les résultats de simulation dynamique sont des estimations du fonctionnement réel du bâtiment. La marge d'erreur présente est tout à fait acceptable et les valeurs estimées restent dans les incertitudes du modèle dynamique initial.

Pour profiter des capacités du modèle polynomial produit dans l'étude du bâtiment il est nécessaire de disposer d'un mode de représentation dynamique qui permette de facilement propager à travers le méta-modèle les choix faits sur les différentes entrées. Pour cela, une interface graphique est réalisée.

4.1 Présentation de l'interface graphique

La Figure IV-6 présente l'affichage configuré pour les huit entrées et quatre sorties. Celle-ci rassemble dans un même affichage l'analyse graphique, l'analyse de

sensibilité et l'analyse d'incertitude. Elle est séparée en deux parties, en bas la sélection des entrées et le résultat du calcul de sensibilité et en haut l'affichage des sorties.

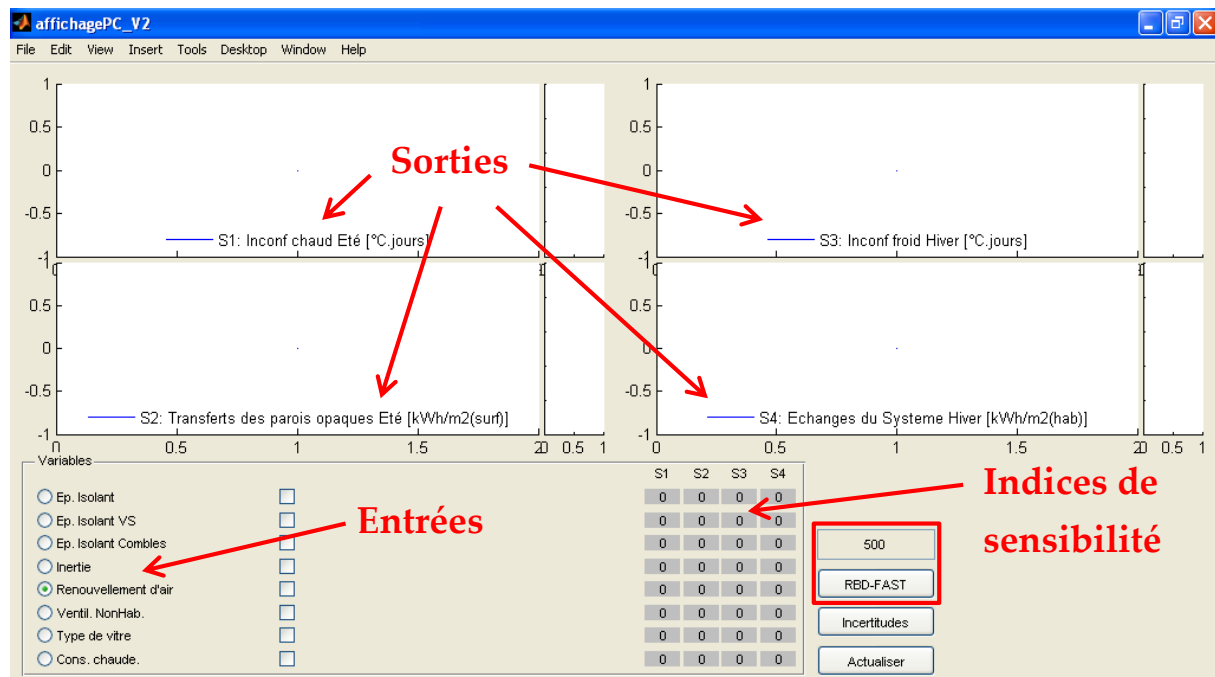


Figure IV-6 : Vue initiale

La liste des entrées présentes est affichée en bas du graphique, devant chacune de celle-ci est placé un bouton radio qui permet de sélectionner l'entrée à afficher sur les graphiques. A droite des entrées, sont présentées des cases à cocher qui offrent la possibilité à l'utilisateur de fixer les variables à des valeurs choisies, mais cette fonctionnalité sera plus amplement détaillée dans la suite.

En fin de ligne est présent un tableau dans lequel seront affichés les indices de sensibilité calculés. La méthode utilisée ainsi le nombre de simulations peuvent être modifiés par l'intermédiaire de la zone de saisie, qui définit le nombre de simulation (500) et le bouton à bascule, qui change la méthode de RBD-FAST à RBD-SOBOL. Pour plus de lisibilité, une coloration est réalisée sur les résultats affichés dans le tableau, les valeurs hautes tendent vers le rouge et les valeurs basses tendent vers le jaune. En cas d'indice très faible ou si le calcul n'a pas été réalisé la couleur de fond d'interface est attribuée. Chaque colonne du tableau représente une sortie numérotée. Ces numéros ainsi que les noms correspondants se retrouvent dans les légendes des graphiques présents dans la partie supérieure de l'affichage. La numérotation se fait de haut en bas puis de gauche à droite.

L'affichage des graphiques reprendra la mise en forme utilisée lors de l'exportation des images. Les résultats seront sous la forme de nuages de points (250) avec une

coloration effectuée sur l'entrée discrète. Dans le cas de l'étude présentée ici, l'entrée discrète est le type de vitrage.

À gauche de chaque sortie sont également présents des graphiques annexes, une représentation de la dispersion de la sortie sous forme de *boxplot* communément appelé boîte à moustache. Ce type de graphique détaillé en Figure IV-7, est utilisé en statistique pour donner une représentation unidimensionnelle des différents quartiles de la distribution. La boîte, définit par sa longueur les valeurs du premier et du troisième quartile (soit 25% et 75%). Celle-ci est séparée par la médiane (50%) et la longueur des moustaches détermine les limites de la distribution. Ces limites correspondent à la valeur haute et basse, mais les moustaches sont contraintes à une longueur maximale égale à 1,5 la longueur de la boîte. Cette longueur correspond dans le cas d'une distribution gaussienne approximativement à 2 et 98 centiles. En cas de dépassement de cette limite, les valeurs extrêmes sont détaillées par des croix. Le dernier élément est un cercle représentant la moyenne de toutes les valeurs présentes.

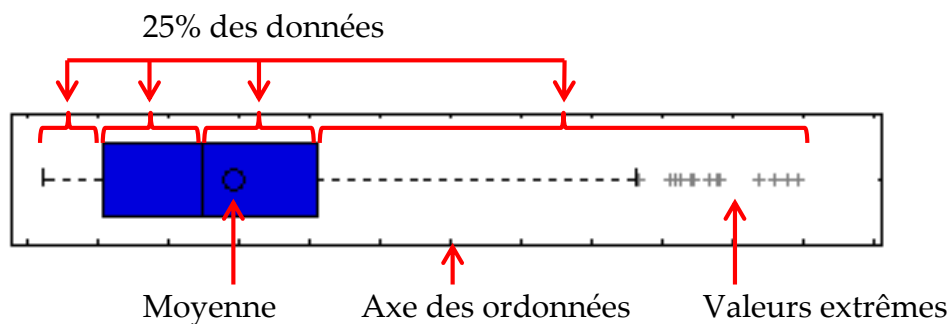


Figure IV-7 : Description d'une boîte à moustache (boxplot)

A chaque appui sur le bouton « Actualiser », les informations relatives à la configuration des entrées sont envoyées au générateur LHS qui réalise un plan de simulation puis au méta-modèle qui réalise les évaluations. Ces résultats alimentent les différents graphiques et sont envoyés vers la méthode d'analyse choisie pour estimer les indices de sensibilité affichés dans le tableau.

4.2 Lecture des données

La Figure IV-8 représente l'interface avec les données du méta-modèle. Celle-ci est configurée avec le renouvellement d'air désigné comme variable de référence. Les points présents sur les graphiques sont organisés selon les valeurs croissantes de cette entrée et montrent des tendances pour 3 des 4 sorties. Le tableau des indices de sensibilité montre alors que cette variable dispose d'effets importants pour ces sorties, mais donne également toutes les autres relations qu'il est possible de

visualiser. On remarque ainsi que la consigne de chauffage est particulièrement influente sur le confort.

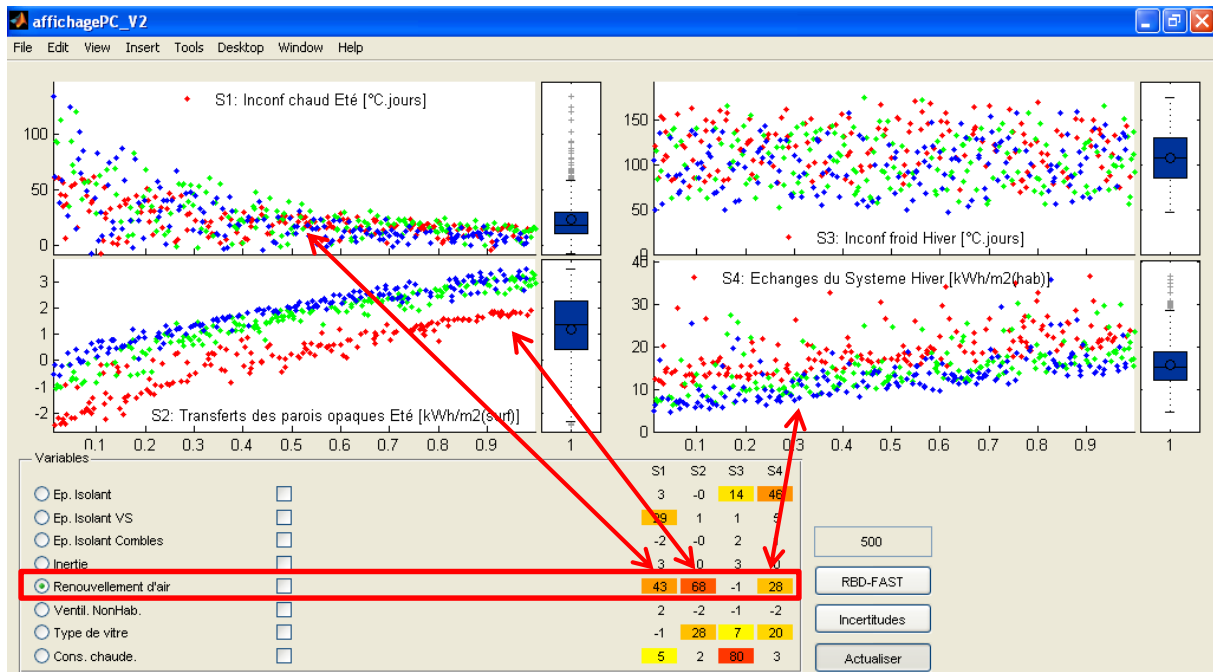


Figure IV-8 : Tendances selon le renouvellement d'air

Le choix d'une autre variable va permettre de modifier l'organisation des graphiques et d'afficher d'autres tendances. La Figure IV-9 montre les résultats organisés selon les valeurs croissantes d'épaisseur d'isolation des murs. Les évaluations réalisées sont sensiblement les mêmes, car les propriétés des variables n'ont pas été modifiées, mais la nouvelle organisation fait apparaître d'autres tendances.

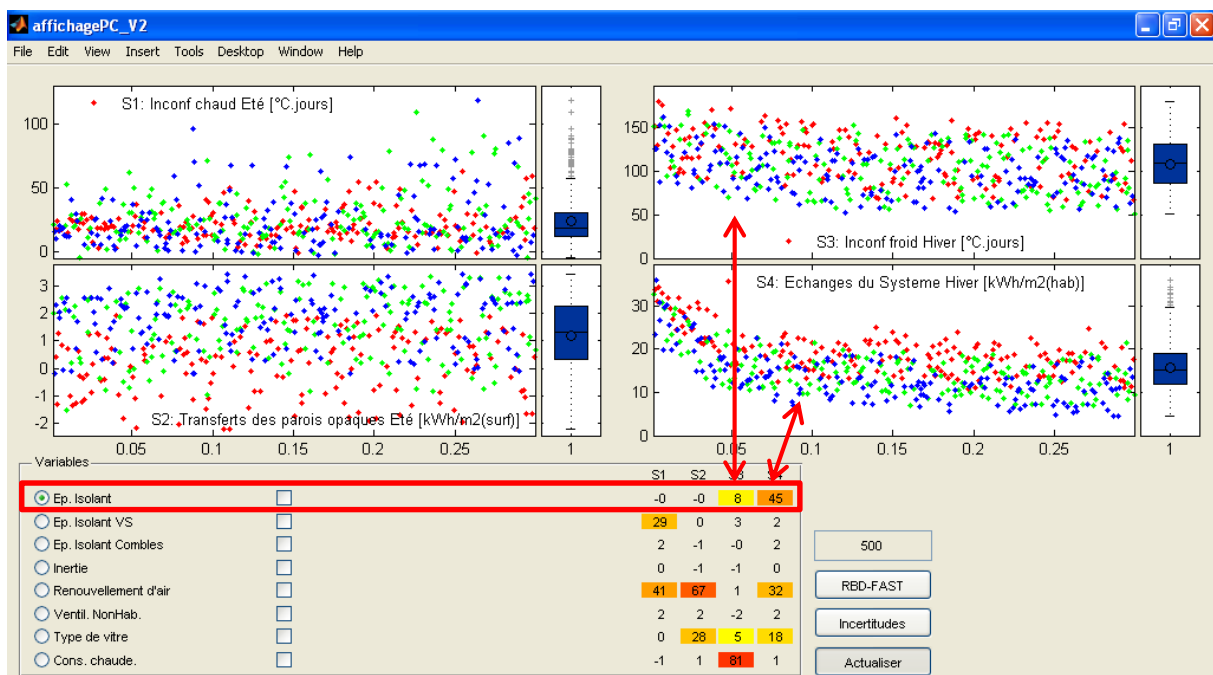


Figure IV-9 : Tendances selon l'épaisseur d'isolation des murs

Les évolutions sont cependant plus dispersées dues à la forte influence du renouvellement d'air qui est maintenant aléatoire. Fixer cette variable va permettre d'obtenir une meilleure évaluation du comportement des autres variables. La Figure IV-10 présente l'interface après avoir fixé le renouvellement d'air. Une barre de défilement apparaît pour permettre de sélectionner la valeur à fixer pour la variable, cette valeur est affichée après la barre. Pour l'instant les valeurs cochées sont fixes et des *NaN* sont affichés dans le tableau, car aucun indice de sensibilité ne peut être déterminé pour une valeur fixe.

Le comportement des sorties est alors plus clair aussi bien au niveau des graphiques que du tableau. On voit donc apparaître le comportement de l'épaisseur d'isolant des murs, mais également l'effet du type de vitrage qui est visible par la coloration réalisée.

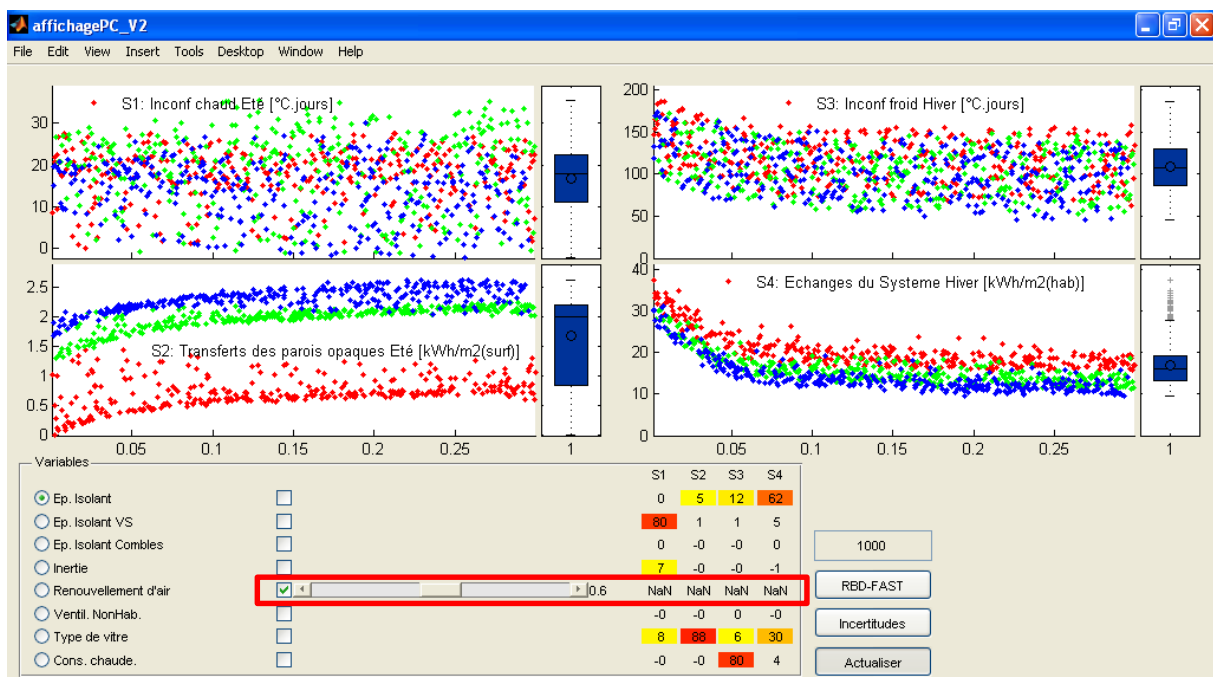


Figure IV-10 : Ajout d'une contrainte

L'étude peut alors se poursuivre, la méthode étant la même que ce qui a été décrit dans le Chapitre III. Le choix des variables se fait à l'aide du tableau et le choix de la valeur en fonction des graphiques. L'utilisateur décide quelle variable fixer en cochant les cases présentes et choisit la valeur adéquate.

Lors du choix de la valeur, l'utilisateur base son choix en fonction du comportement de la variable qu'il étudie, mais les graphiques peuvent présenter une certaine dispersion ce qui rend la lecture des courbes délicate. La Figure IV-11 permet de voir qu'il est possible de fixer temporairement les autres variables influentes afin

d'extraire du méta-modèle le comportement de la seule variable considérée. Le résultat obtenu est alors l'évolution claire de cette variable.

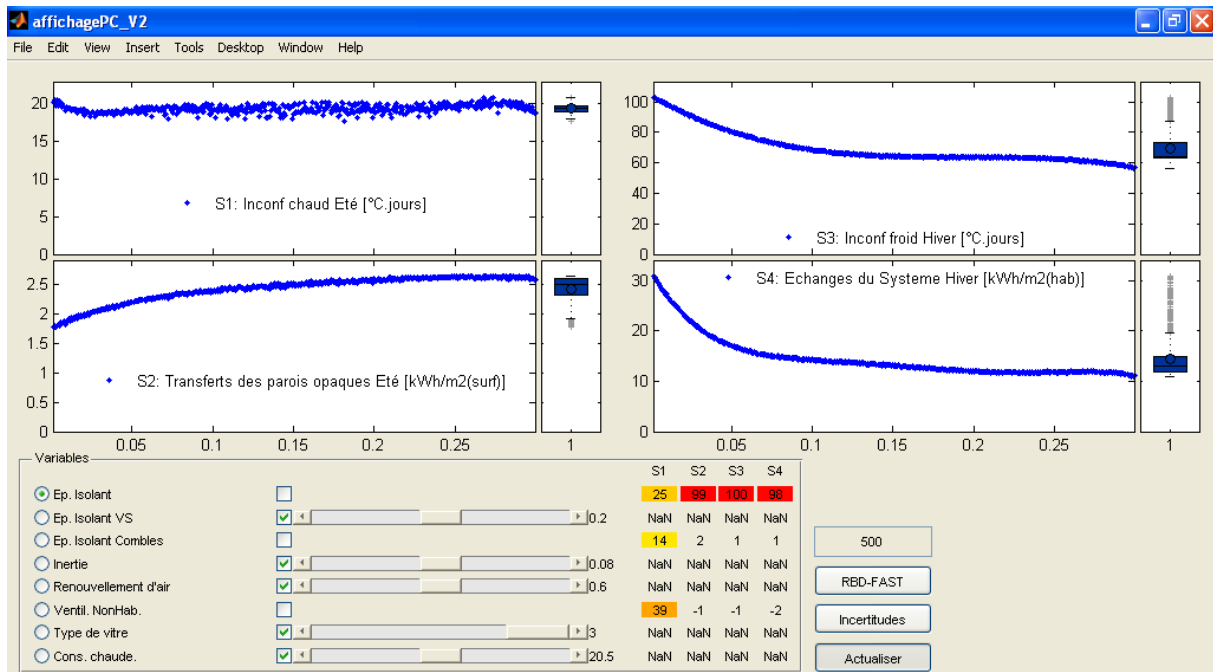


Figure IV-11 : Comportement sans interactions

Cela suppose cependant que toutes les variables sont fixées, or en pratique une incertitude est présente sur la valeur attribuée à chaque variable et parmi elles certaines étaient influentes. La Figure IV-12 montre l'utilisation du bouton à bascule « Incertitudes » qui permet d'activer ou de désactiver les incertitudes sur les entrées sélectionnées. Ainsi, les entrées ne seront plus fixes, mais variables autour des valeurs sélectionnées. La variabilité de chaque entrée étant définie par l'incertitude locale qui lui a précédemment été attribuée (voir Chapitre III :3.3 au Tableau III-9). Les indices de sensibilité peuvent alors être calculés pour toutes les entrées à l'exception des entrées discrètes pour lesquelles l'incertitude locale n'a ici pas de sens.

On réalise ainsi une étude mixte, certaines entrées varient sur tout leur domaine de définition alors que les autres ne varient que de quelques pour cent autour de la solution. Les résultats montrent une forte augmentation de la dispersion pour l'inconfort d'hiver due à l'incertitude et l'influence de la température de consigne. La valeur choisie pour l'épaisseur d'isolation a alors presque autant d'influence qu'une valeur « fixe » vis-à-vis de l'inconfort.

Ce type d'évaluation permet de prendre compte l'importance des perturbations et de comparer les incertitudes avec le choix réalisé. La valeur à attribuer à la variable

considérée ne doit ainsi pas tenir compte de ces sorties fortement perturbées, car d'autres entrées auront des impacts beaucoup plus significatifs.

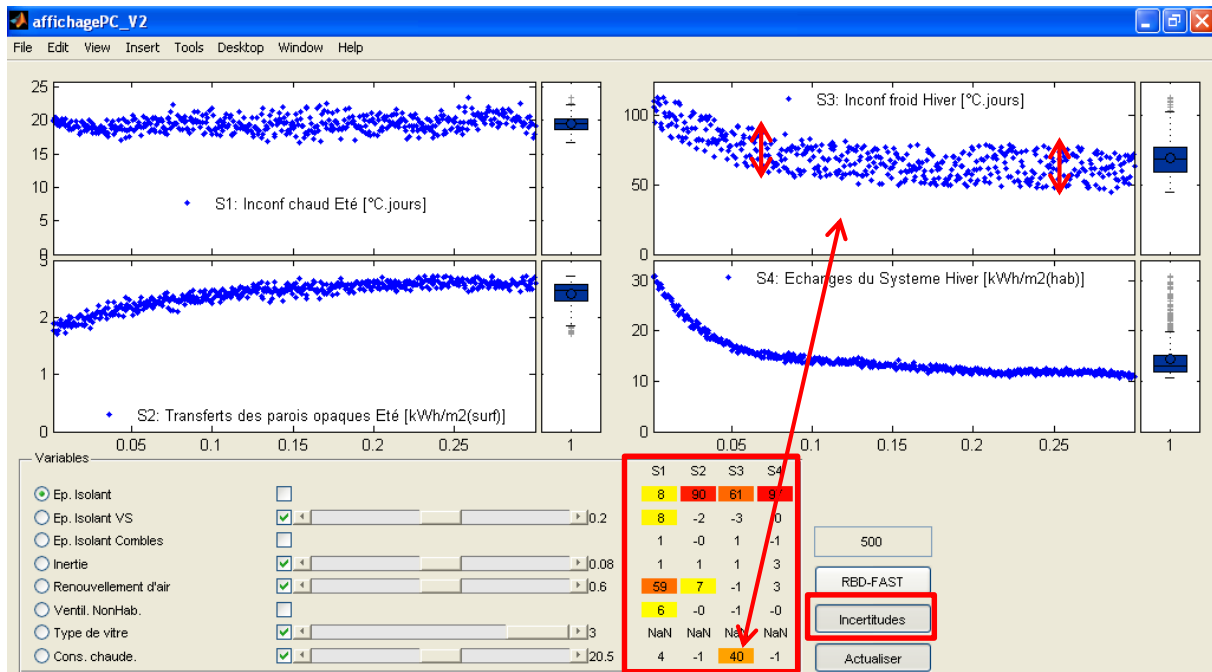


Figure IV-12 : Présentation de l'interface graphique – prise en compte des incertitudes dans la recherche

Une fois toutes les entrées définies, il est ainsi possible de passer directement de la recherche de solution à l'analyse d'incertitude. La Figure IV-13 montre les résultats d'analyse d'incertitude. Les graphiques donnent les moyennes et variances attendues vis-à-vis des incertitudes définies et le tableau identifie les entrées qui produisent cette incertitude. Les résultats affichent des valeurs proches de celles obtenues lors de l'étude directe avec le modèle initial et qui ont été présentées en Chapitre III :3.3 au Tableau III-10.

On a alors une estimation robuste des solutions de réhabilitation.

Cela n'a pas été précisé, mais toutes ces représentations du programme ont été limitées pour des raisons de clarté. Celles-ci ont été réalisées sur un écran de petite taille –résolution de 1024x600. L'affichage de l'interface s'adapte en fonction de la taille de l'écran. La Figure IV-14 présente l'interface affichée sur un écran de bureau – avec une résolution de 1920x1080. Une étude plus complète peut ainsi être menée avec un nombre plus important de sorties ou de variables sans altérer la lisibilité des résultats.

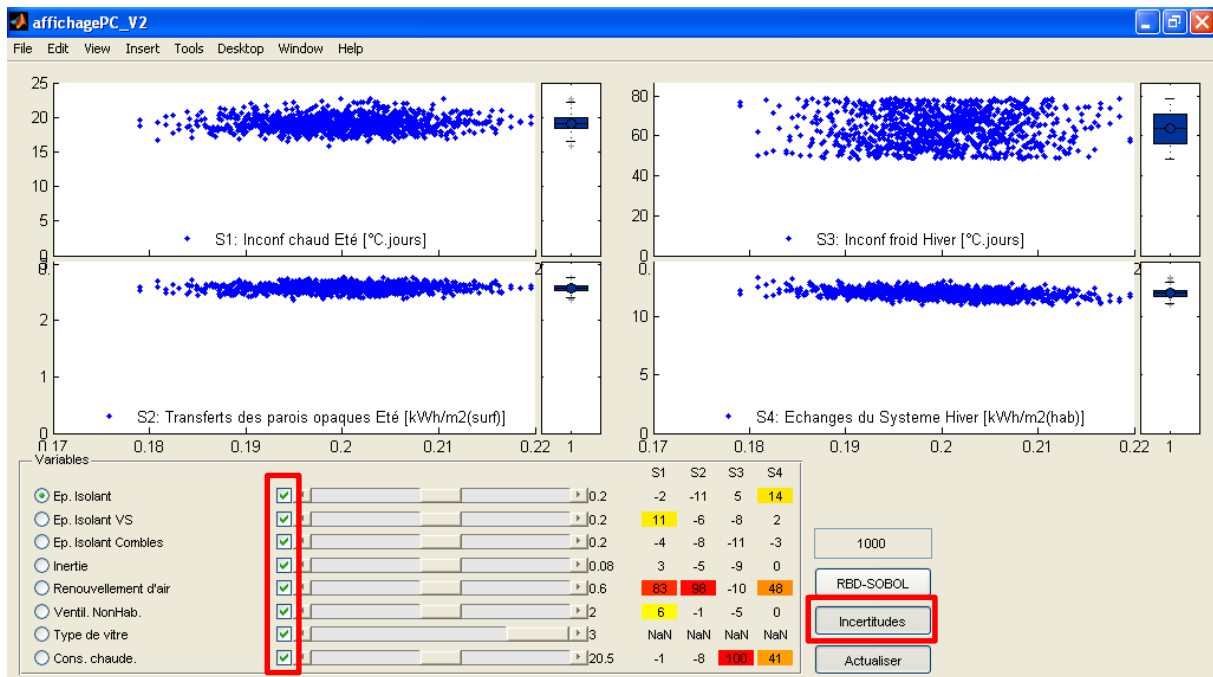


Figure IV-13 : Incertitudes sur la solution finale

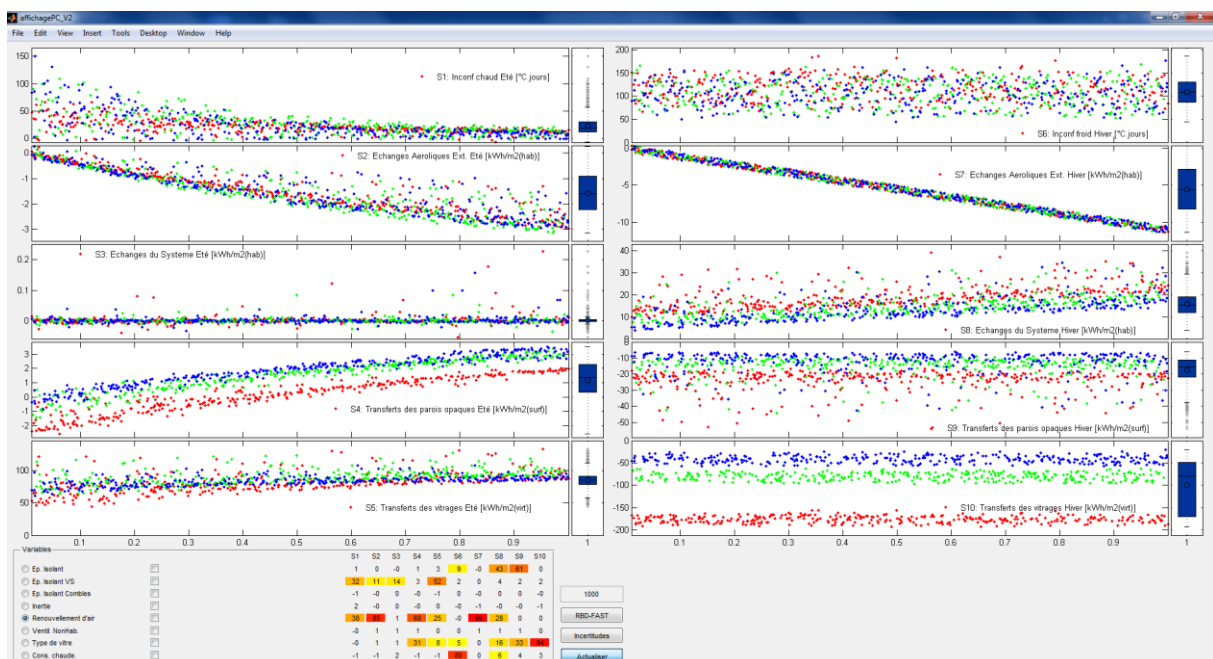


Figure IV-14 : Vue classique avec 12 sorties

5 Le méta-modèle une réponse aux contraintes de l'étude thermique

Dans ce dernier chapitre les analyses de sensibilité et le programme de simulation ont été utilisés afin de produire un affichage qui permet d'appliquer la démarche de recherche de solution qui a été pensée dans le Chapitre III.

Il a été décidé d'aller à l'opposé des outils d'optimisation, les simulations ne servent plus à rechercher une solution, mais à analyser complètement le comportement. Les sorties ont montré des évolutions relativement simples et monotones vis-à-vis de chaque entrée. Le criblage de l'espace réalisé, permet alors d'alimenter un modèle polynomial qui peut à son tour donner une estimation des réponses afin d'évaluer le comportement général des sorties malgré la présence d'interactions et de légères discontinuités. Ce nouveau modèle permet d'obtenir une approximation du modèle initial quel que soit la combinaison de valeur d'entrée demandée à condition de rester sur l'espace défini pour le criblage.

L'analyse est ainsi optimisée, car toutes les informations utiles ont été extraites des simulations dynamiques, ce qui permet ensuite de réaliser un travail plus poussé dans l'étude du bâtiment. Le coût de simulation n'est plus une contrainte.

La mise en place d'une interface graphique apporte alors beaucoup d'informations sur le comportement de chaque entrée vis-à-vis des performances observées, car en plus de l'identification des relations influentes elle permet d'identifier l'impact des incertitudes. Les différents choix d'optimisation peuvent alors se faire en comparant les variations produites en sortie par les différentes entrées, en fonction des incertitudes en sortie engendrées par la méconnaissance de ces mêmes entrées.

Les variables sont modifiées en connaissant exactement l'impact qu'elles ont sur le modèle, une vision globale du bâtiment est proposée avec plus de transparence sur les entrées. Le choix de la solution est le résultat d'un compromis fait au cas par cas qui peut alors être maîtrisé tout en prenant en compte les incertitudes.

CONCLUSION

La réhabilitation du parc immobilier existant est un enjeu actuel majeur pour tendre vers une réduction importante de nos consommations énergétiques. Les technologies actuelles permettent de construire des bâtiments à basse consommation, mais sont difficilement appliquées à l'existant qui subit une mutation lente et peu encadrée. L'amélioration des performances nécessite alors une recherche de solution de réhabilitation et passe nécessairement par l'utilisation de logiciel de simulation. Ces logiciels permettent d'estimer le comportement fin du bâtiment avant et après la réhabilitation et peuvent ainsi évaluer avant travaux les meilleures pistes d'amélioration du bâti. Ces outils puissants sont cependant opaques et complexes de par le nombre important d'entrées que l'on doit renseigner et de par les interactions de phénomènes physiques au sein du bâtiment à différentes échelles et niveau. Pour permettre un choix optimisé de solution de réhabilitation, il devient nécessaire d'utiliser des méthodes numériques pour évaluer les différentes combinaisons. Cependant, ces solutions doivent pouvoir être mises en œuvre et répondent nécessairement à des contraintes multiples qui sont source de compromis. L'intégration au sein de l'optimisation est alors délicate et il est envisagé d'aborder de nouvelles approches dans l'utilisation des simulations numériques pour la recherche de solution.

Pour répondre à cette problématique, certaines méthodes issues du domaine de l'analyse de sensibilité sont explorées. Seules les méthodes adaptées aux modèles complexes à haute dimension –tels que les modèles de thermique du bâtiment– sont prises en compte. Le formalisme de telles méthodes statistiques est décrit en Chapitre II. Seulement, le coût de simulation est une contrainte importante et qui a été considérée en sélectionnant des méthodes performantes pour l'estimation de l'effet des principales variables d'entrée avec peu d'appels au modèle. Il s'agit entre autres de RBD-FAST, qui par une analyse fréquentielle permet d'estimer les entrées les plus influentes sur les sorties du modèle comme le taux d'inconfort ou les besoins énergétiques. La bonne mise en œuvre de ces méthodes a été assurée avec la mise en place de méthodes d'échantillonnage reconnues qui génèrent des variations d'entrée correctes. Les résultats sont également vérifiés avec une technique de bootstrap pour évaluer l'erreur ou par comparaison avec d'autres techniques telle que RBD-SOBOL.

L'application de ces méthodes a été réalisée au Chapitre III via la mise en place d'un programme. Ce programme permet une utilisation quasi-automatique des analyses de sensibilité allant de la création des fichiers de simulation jusqu'à la présentation d'indices de sensibilité et de graphiques associés. Cela a permis de réaliser deux analyses, d'une part sur un cas test simple et ensuite sur une des maisons individuelles de la plateforme expérimentale nommée INCAS située à l'INES à Chambéry. Les résultats montrent les possibilités de telles méthodes statistiques, même dans le cas complexe d'un bâtiment entier, et permettent de différencier les saisons été et hiver. Cette étude identifie les principales variables d'entrée régissant le comportement des sorties. Les entrées ont des influences très différentes en fonction de la saison et des sorties considérées permettant ainsi de réaliser une sélection progressive et multicritère des valeurs d'entrées. Une représentation graphique de type *scatterplot* permet une visualisation de l'effet des entrées influentes sur les sorties, sur l'ensemble de la plage de variation de l'entrée, ce qui permet d'estimer rapidement la zone d'effet et la valeur optimale de l'entrée par rapport à la sortie considérée. Une étude croisée sur plusieurs sorties et entrée amène une connaissance du modèle poussée pour permettre le choix des interventions énergétiques. Ainsi à partir d'un modèle thermique de bâtiment existant, la méthode RBD FAST sonde la structure du modèle et en détermine les entrées les plus influentes sur lesquelles une action peut être envisagée dans le cas d'une opération de réhabilitation. Un autre avantage de ces méthodes est qu'elles sont initialement destinées à l'évaluation des incertitudes. La solution retenue peut alors être étudiée en vue d'assurer une garantie de performance ou de façon plus sommaire d'évaluer une plage d'incertitude vis-à-vis des investissements devant être réalisés.

Cependant, le coût de simulation d'une telle étude rend l'application de cette méthodologie délicate face aux contraintes industrielles. L'augmentation du nombre de sorties et l'aspect multicritère amènent une complexité supplémentaire qui nécessite une analyse plus fine que les cas présentés. Or à chaque fois qu'une entrée est fixée ou perturbée, une nouvelle étude de l'ordre de la dizaine d'heures CPU doit être réalisée pour évaluer un nouveau jeu de simulations généralement très semblables aux simulations déjà obtenues. Le coût de simulation augmente donc rapidement et pour répondre à ce verrou technologique la piste des méta-modèles est explorée en Chapitre IV. La technique du chaos polynomial est choisie afin de produire un modèle simplifié au modèle complexe de bâtiment et de s'assurer de la bonne évaluation des prochains indices de sensibilité basés sur ce modèle. Avoir recours à un modèle intermédiaire est une méthode couramment employée pour étudier les modèles qui nécessitent un coût de calcul important. L'emploi de cette

technique permet ainsi de s'assurer d'une estimation fiable des indices et permet de réaliser d'autres évaluations de façon quasi-instantanée.

Ce nouveau modèle polynomial est alors intégré dans une interface graphique afin de produire un cadre propice pour l'étude du modèle et la recherche de solutions. Celle-ci s'adapte en fonction du nombre de variables et du nombre de sorties étudiées et fournit une vue synthétique du comportement afin de réaliser la démarche de recherche de solution précédemment décrite. La mise à jour des graphiques se fait par l'intermédiaire du modèle polynomial et donc de façon quasi-instantanée. Les évaluations réalisées peuvent alors porter sur l'ensemble du domaine, sur une solution particulière ou sur une définition alternative entre les deux, c'est-à-dire que toutes les variables explorent initialement l'intégralité de leur domaine de définition, mais qu'il est possible de leur attribuer une valeur particulière choisie par l'utilisateur. Ces valeurs particulières correspondent aux propriétés idéales retenues et peuvent, au choix de l'utilisateur, disposer tout de même des incertitudes relatives à la méconnaissance qui sera présente lors de la mise en œuvre. A chaque mise à jour des graphiques une étude de sensibilité est réalisée avec la méthode RBD-FAST afin d'orienter l'utilisateur dans ses prochains choix et lui permettre ainsi de choisir les valeurs des entrées par ordre d'influence. En fin d'analyse, les résultats obtenus avec toutes les entrées déterminées et les incertitudes activées indiquent alors une estimation de l'erreur attendue due aux incertitudes définies. Les indices de sensibilité présents indiquent alors quelles variables produisent cet écart afin de permettre un meilleur contrôle des celles-ci.

Finalement, un outil complet de recherche-validation est ainsi généré afin de répondre aux contraintes méthodologiques soulignées en Chapitre I. La démarche se base sur la volonté d'apporter plus de transparence sur la simulation thermique dynamique pour obtenir rapidement des informations clés et éviter les erreurs d'utilisateur. Le concepteur est au centre de la démarche, il est amené à analyser le comportement global de son modèle et peut ainsi détecter les erreurs lorsque celles-ci produisent des comportements non-attendus. De plus, il est l'acteur du choix de solution en fonction des différentes informations obtenues et réalise directement l'arbitrage entre les différents points de l'évaluation multicritère. L'approche boîte noire de l'optimisation est troquée pour une approche transparente de visualisation du squelette du modèle. C'est cette notion d'assistance au concepteur pour le gain en transparence qui semble être la plus indiquée pour tendre vers les exigences de plus en plus ambitieuses de réduction énergétiques tout en prenant en compte les contraintes de projet.

PERSPECTIVES

L'exploration des capacités du méta-modèle au cours de ces travaux a été faite à partir d'un modèle de bâtiment validé et calibré. Afin d'envisager l'application de la méthode développée dans ces travaux de thèse, il est important d'employer des méthodes de mesure in situ fiables pour permettre la calibration de modèles de bâtiments occupés dans le parc immobilier existant ce qui constitue une première voie d'exploitation des travaux réalisés.

La validation de la solution retenue est également un point important qui conduira à la garantie de performance en site occupé. Cependant l'estimation des entrées et des incertitudes associées reste délicate dans le cas d'une réhabilitation où le diagnostic initial est toujours un verrou difficile à lever et un travail important sur la prise en compte des incertitudes reste nécessaire.

Une autre difficulté dans l'estimation des consommations est liée aux paramètres non maîtrisables tels que les scénarios d'occupation ou les conditions climatiques très incertaines. Il s'agira donc là aussi de fournir une évaluation fiable de l'écart attendu au niveau des sorties du modèle afin d'envisager une garantie de performance qui jouera un rôle prépondérant dans la réduction des consommations.

On peut également envisager la mise en place d'un algorithme d'optimisation pour proposer à moindre coût une solution de base ou pour accompagner le concepteur dans sa démarche tout au long de l'opération de réhabilitation.

Bien qu'illustrés dans le contexte de réhabilitation les outils statistiques présentés dans ces travaux ont été sélectionnés de par leur caractère générique et non intrusif. Ainsi ils peuvent également répondre à des problématiques différentes telles que l'optimisation de bâtiments neufs ou la justification de solutions technologiques innovantes.

BIBLIOGRAPHIE

- ADEME. « Centre de ressources pour les Plans Climat-Energie Territoriaux », disponible sur: <http://www.pcet-ademe.fr/domaines-actions/batiments/contexte-et-enjeux> (consulté le 26 septembre 2014).
- Archer, G.E.B., Saltelli, A. et Sobol, I.M. (1997), « Sensitivity measures, anova-like Techniques and the use of bootstrap », *Journal of Statistical Computation and Simulation*, Vol. 58 No. 2, p. 99-120.
- ASHRAE. (2001), « Modeling two- and three-dimensional heat transfer through composite wall and roof assemblies in transient energy simulation programs », *Research Project 1145-TRP, Technical Committee 4.7*, disponible sur: http://web.ornl.gov/sci/roofs+walls/research/detailed_papers/whole_bldg/ (consulté le 26 septembre 2014).
- Bhandari, M., Shrestha, S. et New, J. (2012), « Evaluation of weather datasets for building energy simulation », *Energy and Buildings*, Elsevier B.V., Vol. 49, p. 109-118.
- Blatman, G. (2009), *Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis*, Université Blaise Pascal - Clermont II.
- Blatman, G. et Sudret, B. (2009), « Efficient global sensitivity analysis of computer simulation models using an adaptive Least Angle Regression scheme », *41èmes Journées de Statistique*, p. 1-6.
- Blatman, G. et Sudret, B. (2010), « An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis », *Probabilistic Engineering Mechanics*, Elsevier Ltd, Vol. 25 No. 2, p. 183-197.
- Bluyssen, P.M. (2000), « EPIQR and IEQ: indoor environment quality in European apartment buildings », *Energy and Buildings*, Vol. 31 No. 2, p. 103-110.
- Boerstra, A. (2006), *The adaptive thermal comfort criterion in the new EPBD IEQ standard*, disponible sur: <http://nceub.commoncense.info/uploads/Boerstra.pdf> (consulté le 26 septembre 2014).
- Booth, K. et Cox, D. (1962), « Some systematic supersaturated designs », *Technometrics*, Vol. 4 No. 4, p. 489-495.

- Butler, J., Jia, J. et Dyer, J. (1997), « Theory and Methodology Simulation techniques for the sensitivity analysis of multi-criteria decision models », *European Journal of Operational Research*, No. 103, p. 531-546.
- Cameron, R.H. et Martin, W.T. (1947), « The Orthogonal Development of Non-Linear Functionals in Series of Fourier-Hermite Functionals », *The Annals of Mathematics*, Vol. 48 No. 2, p. 385.
- Carlucci, S. et Pagliano, L. (2012), « A review of indices for the long-term evaluation of the general thermal comfort conditions in buildings », *Energy and Buildings*, Elsevier B.V., Vol. 53, p. 194-205.
- Castaigns, W. (2007), *Analyse de sensibilité et estimation de paramètres pour la modélisation hydrologique: potentiel et limitations des méthodes variationnelles*, Université Joseph Fourier - Grenoble I.
- CCNUCC. (1994), *Convention-Cadre des Nations Unies sur les Changements Climatiques (CCNUCC)*, disponible sur: http://unfccc.int/portal_francofone/essential_background/convention/items/3270.php (consulté le 26 septembre 2014).
- Chastaing, G. (2013), « Indices de Sobol généralisés pour variables dépendantes ».
- Cheng, C. et Pouffary, S. (2008), *The Kyoto protocol, the clean development mechanism and the building and construction sector – A Report for the UNEP Sustainable Buildings and Construction Initiative*, (United Nations Environment Programme, Éd.), Paris.
- Conseil d'Architecture d'Urbanisme et de l'Environnement. « Réglementation thermique et labels », disponible sur: <http://www.caue44.com/IMG/pdf/02reglementation.pdf> (consulté le 26 septembre 2014).
- Cukier, R.I. (1973), « Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory », *The Journal of Chemical Physics*, Vol. 59 No. 8, p. 3873.
- Cukier, R.I., Schaibly, J.H. et Shuler, K.E. (1975), « Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. III. Analysis of the approximations », *The Journal of Chemical Physics*, Vol. 63 No. 3, p. 1140.
- Dear, R. De et Brager, G. (1998), « Developing an adaptive model of thermal comfort and preference », *ASHRAE Transactions*, Vol. 104 No. 1, p. 145-167.

- De Dear, R.J., Brager, G.S. et Cooper, D. (1997), « Developing an adaptive model of thermal comfort and preference », ASHRAE RP-884 - Final Report.
- Directive 2010/31/UE. (2010), « Directive 2010/31/UE sur la performance énergétique des bâtiments (refonte) », *Journal officiel de l'Union européenne*, p. 13-35.
- DRAFT prEN 15251. (2005), « Criteria for the Indoor Environment including thermal, indoor air quality, light and noise », *Comité Européen de Normalisation*, No. E, p. 1-44.
- Dubreuil, S., Berveiller, M., Petitjean, F. et Salaün, M. (2014), « Construction of bootstrap confidence intervals on sensitivity indices computed by polynomial chaos expansion », *Reliability Engineering & System Safety*, Vol. 121, p. 263-275.
- EC Adhoc Industrial Advisory Group. (2013), *Energy efficient buildings Public-Private Partnership (PPP): Multi annual roadmap and longer term strategy*, European Commission, DG Research.
- Efron, B. (1992), « Jackknife-after-bootstrap standard errors and influence functions », *Journal of the Royal Statistical Society*, Vol. 54 No. 1, p. 83-127.
- Efron, B., Hastie, T., Johnstone, I. et Tibshirani, R. (2004), « Least angle regression », *The Annals of Statistics*, Vol. 32 No. 2, p. 407-499.
- EN 15217. (2008), *Performance énergétique des bâtiments - Méthodes d'expression de la performance énergétique et de certification énergétique des bâtiments*, p. P 50-780.
- EN 15251. (2007), *Critères d'ambiance intérieure pour la conception et évaluation de la performance énergétique des bâtiments couvrant la qualité de l'air intérieur , la thermique , l'éclairage et l'acoustique*, p. E 51-762.
- EN ISO 7730. (2006), *Ergonomie des ambiances thermiques Détermination analytique et interprétation du confort par des critères de confort thermique local*, p. X 35-203.
- Fang, K.-T., Li, R. et Sudjianto, A. (2005), *Design and Modeling for Computer Experiments*, (Chapman and Hall/CRC Computer Science & Data Analysis, Éd.).
- FD X 50-171. (2000), *Système de management de la qualité - Indicateurs et tableaux de bord*, p. X 50-171.
- Franco, J. (2008), *Planification d'expériences numériques en phase exploratoire pour la simulation des phénomènes complexes*, Ecole Nationale Supérieure des Mines de Saint-Etienne.

- Fumo, N., Mago, P. et Luck, R. (2010), « Methodology to estimate building energy consumption using EnergyPlus Benchmark Models », *Energy and Buildings*, Elsevier B.V., Vol. 42 No. 12, p. 2331-2337.
- Garcia-Cabrejo, O. et Valocchi, A. (2014), « Global sensitivity analysis for multivariate output using polynomial chaos expansion », *Reliability Engineering & System Safety*, Elsevier, Vol. 126, p. 25-36.
- GDR MASCOT-NUM. « Goupement de Recherche Méthode d'Analyse Stochastique pour les COdes et Traitements NUMériques », disponible sur: <http://www.gdr-mascotnum.fr> (consulté le 26 septembre 2014).
- Genre, J.-L., Flourentzou, F. et Stockli, T. (2000), « Building refurbishment: habitat upgrading », *Energy and Buildings*, Vol. 31 No. 2, p. 155-157.
- Goffart, J. (2013), *Impact de la variabilité des données météo sur une maison basse consommation. Application des analyses de sensibilité pour des entrées temporelles corrélées.*, Université de Savoie.
- Hamdy, M., Hasan, A. et Siren, K. (2011), « Impact of adaptive thermal comfort criteria on building energy use and cooling equipment size using a multi-objective optimization scheme », *Energy and Buildings*, Elsevier B.V., Vol. 43 No. 9, p. 2055-2067.
- Helton, J.C. (1993), « Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal », *Reliability Engineering & System Safety*, Vol. 42 No. 2-3, p. 327-367.
- Helton, J.C. et Davis, F.J. (2002), « Illustration of sampling-based methods for uncertainty and sensitivity analysis. », *Risk analysis : an official publication of the Society for Risk Analysis*, Vol. 22 No. 3, p. 591-622.
- Helton, J.C. et Davis, F.J. (2003), « Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems », *Reliability Engineering & System Safety*, Vol. 81 No. 1, p. 23-69.
- Homma, T. et Saltelli, A. (1996), « Importance measures in global sensitivity analysis of nonlinear models », *Reliability Engineering & System Safety*, Vol. 52, p. 1-17.
- Hopfe, C.J. et Hensen, J.L.M. (2011), « Uncertainty analysis in building performance simulation for design support », *Energy and Buildings*, Elsevier B.V., Vol. 43 No. 10, p. 2798-2805.

- Iman, R.L. et Hora, S.C. (1990), « A robust measure of uncertainty importance for use in fault tree system analysis », *Risk analysis*, No. 3, p. 401-406.
- Iooss, B. (2011), « Revue sur l'analyse de sensibilité globale de modèles numériques », *Journal de la Société Française de Statistique*, p. 1-23.
- IPCC. (2013), *Climate Change 2013: The Physical Science Basis*, Intergovernmental Panel on Climate Change (IPCC).
- Jaffal, I., Inard, C. et Ghiaus, C. (2009), « Fast method to predict building heating demand based on the design of experiments », *Energy and Buildings*, Vol. 41 No. 6, p. 669-677.
- Janssen, H. (2013), « Monte-Carlo based uncertainty analysis: Sampling efficiency and sampling convergence », *Reliability Engineering & System Safety*, Elsevier, Vol. 109, p. 123-132.
- Johnson, M., Moore, L. et Ylvisaker, D. (1990), « Minimax and maximin distance designs », *Journal of Statistical Planning and Inference*, Vol. 26 No. 2, p. 131-148.
- Joint Research Centre website. « Sensitivity Analysis routines », disponible sur: <http://ipsc.jrc.ec.europa.eu/?id=756> (consulté le 26 septembre 2014).
- Krzykacz, B. (1990), *SAMOS: a computer program for the derivation of empirical sensitivity measures of results from large computer models*. Technical Report GRS-A-1700, Garching, Gesellschaft fuer Reaktor Sicherheit (GRS).
- Kucherenko, S., Tarantola, S. et Annoni, P. (2012), « Estimation of global sensitivity indices for models with dependent variables », *Computer Physics Communications*, Vol. 183, p. 937-946.
- Kuipers, L. et Niederreiter, H. (2006), *Uniform distribution of sequences*, Dover, p. 416.
- Lam, J.C., Wan, K.K.W. et Yang, L. (2008), « Sensitivity analysis and energy conservation measures implications », *Energy Conversion and Management*, Vol. 49 No. 11, p. 3170-3177.
- Larceneux, F. (2010), « Les facteurs psycho-sociologiques de l'achat d'un logement », *Marketing de l'immobilier*, p. 1-43.
- Liu, J.S. (2001), *Monte Carlo strategies in scientific computing*, Springer., p. 343.

- Mara, T.A. et Rakoto Joseph, O. (2008), « Comparison of some efficient methods to evaluate the main effect of computer model factors », *Journal of Statistical Computation and Simulation*, Vol. 78 No. 2, p. 167-178.
- Mara, T.A. et Tarantola, S. (2008), « Application of global sensitivity analysis of model output to building thermal simulations », *Building Simulation*, Vol. 1 No. 4, p. 290-302.
- Marseguerra, M., Zio, E. et Podofillini, L. (2003), « Model parameters estimation and sensitivity by genetic algorithms », *Annals of Nuclear Energy*, Vol. 30 No. 14, p. 1437-1456.
- McKay, M. (1995), « Evaluating prediction uncertainty », *Statistics Group, Los Alamos National Laboratory*, Vol. NUREG/CR-6 No. LA-12915-MS.
- McKay, M.D.M., Beckman, R.J.R. et Conover, W.J.W. (1979), « A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code », *Technometrics*, Vol. 21 No. 2, p. 239-245.
- Merheb, R. (2013), *Fiabilité des outils de prévision du comportement des systèmes thermiques complexes*, Université de Bordeaux 1.
- Ministère Français de l'Ecologie, du D. durable et de l'Energie. (2013), *Plan commenté du projet de loi de programmation sur la transition énergétique*, disponible sur: http://www.developpement-durable.gouv.fr/IMG/pdf/2013-12-10_-_Plan_commente_PJL_Programmation_TE.pdf (consulté le 26 septembre 2014).
- Monahan, J. et Powell, J.C. (2011), « An embodied carbon and energy analysis of modern methods of construction in housing: A case study using a lifecycle assessment framework », *Energy and Buildings*, Elsevier B.V., Vol. 43 No. 1, p. 179-188.
- Moran, F., Natarajan, S. et Nikolopoulou, M. (2012), « Developing a database of energy use for historic dwellings in Bath, UK », *Energy and Buildings*, Elsevier B.V., Vol. 55, p. 218-226.
- Morris, M. (1991), « Factorial sampling plans for preliminary computational experiments », *Technometrics*, Vol. 33 No. 2, p. 161-174.
- Moujalled, B. (2007), *Modélisation dynamique du confort thermique dans les bâtiments naturellement ventilés*, Institut des Sciences Appliquées de Lyon.

- Nabil, A. et Mardaljevic, J. (2006), « Useful daylight illuminances: A replacement for daylight factors », *Energy and Buildings*, Vol. 38 No. 7, p. 905-913.
- Nagata, A. (2005), « A simple method to incorporate thermal bridge effects into dynamic heat load calculation programs », *International IBPSA Conference*, Montréal, p. 817-822.
- Nicol, J.F. et Humphreys, M.A. (2002), « Adaptive thermal comfort and sustainable thermal standards for buildings », *Energy and Buildings*, Vol. 34, p. 563-572.
- Pohekar, S.D. et Ramachandran, M. (2004), « Application of multi-criteria decision making to sustainable energy planning - A review », *Renewable and Sustainable Energy Reviews*, Vol. 8 No. 4, p. 365-381.
- Rabouille, M., Wurtz, E. et Perrotin, P. (2013), « Analysis of dynamic thermal simulation for refurbishment », *Thirteenth International IBPSA ...*, Chambéry, p. 358-365.
- Rabouille, M., Wurtz, E. et Perrotin, P. (2014), « Outil d'aide à la conception: Automatisation de l'analyse de sensibilité pour EnergyPlus », *IBPSA France*, Arras.
- Rocquigny, E. de, Devictor, N. et Tarantola, S. (2008), *Uncertainty in industrial practice: a guide to quantitative uncertainty management*, *Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki*, Wiley., p. 364.
- Sallaberry, C.J. et Helton, J.C. (2006), « An Introduction to Complete Variance Decomposition », *IMAC: Conference & Exposition on Structural Dynamics 24th*.
- Saltelli, A. (2002), « Making best use of model evaluations to compute sensitivity indices », *Computer Physics Communications*, Vol. 145 No. 2, p. 280-297.
- Saltelli, A., Andres, T. et Homma, T. (1993), « Sensitivity analysis of model output », *Computational Statistics & Data Analysis*, Vol. 15 No. 2, p. 211-238.
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. et Tarantola, S. (2010), « Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index », *Computer Physics Communications*, Elsevier B.V., Vol. 181 No. 2, p. 259-270.

- Saltelli, A. et Bolado, R. (1998), « An alternative way to compute Fourier amplitude sensitivity test (FAST) », *Computational Statistics & Data Analysis*, Vol. 26, p. 445-460.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., et al. (2007), *Global Sensitivity Analysis. The Primer*, John Wiley & Sons, Ltd, Chichester, UK, doi:10.1002/9780470725184.
- Saltelli, A., Tarantola, S., Campolongo, F. et Ratto, M. (2004), *Sensitivity analysis in practice: a guide to assessing scientific models*, Wiley., p. 232.
- Saltelli, A., Tarantola, S. et Chan, K.P.-S. (1999), « A quantitative model-independent method for global sensitivity analysis of model output », *Technometrics*, American Society for Quality and the American Statistical Association, Vol. 41 No. 1, p. 39-56.
- Satterthwaite, F.E. (1959), « Random Balance Experimentation », *Technometrics*, Vol. 1 No. 2, p. 111-137.
- Schaibly, J.H. (1973), « Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. II Applications », *The Journal of Chemical Physics*, Vol. 59 No. 8, p. 3879.
- Shukhman, B. (1994), « Generation of quasi-random (LP τ) vectors for parallel computation », *Computer Physics Communications*, Vol. 78 No. 3, p. 279-286.
- Sicurella, F., Evola, G. et Wurtz, E. (2011), « A statistical approach for the evaluation of thermal and visual comfort in free-running buildings », *Energy and Buildings*, Elsevier B.V., Vol. 47, p. 402-410.
- Sobol', I.M., Tarantola, S., Gatelli, D., Kucherenko, S.S. et Mauntz, W. (2007), « Estimating the approximation error when fixing unessential factors in global sensitivity analysis », *Reliability Engineering & System Safety*, Vol. 92 No. 7, p. 957-960.
- Sobol', I. (1990), « Sensitivity estimates for nonlinear mathematical models, *Matematicheskoe Modelirovanie*, 2, 112- 118 (in Russian), translated in *Mathematical Modelling and Computational Experiments*, 1, 407-414 (1993). »
- Soize, C. et Ghanem, R. (2004), « Physical Systems with Random Uncertainties: Chaos Representations with Arbitrary Probability Measure », *SIAM Journal on Scientific Computing*, Vol. 26 No. 2, p. 395-410.

- Spitz, C. (2012), *Analyse de la fiabilité des outils de simulation et des incertitudes de métrologie appliquée à l'efficacité énergétique des bâtiments*, Université de Savoie.
- Spitz, C., Mora, L., Wurtz, E. et Jay, A. (2012), « Practical application of uncertainty analysis and sensitivity analysis on an experimental house », *Energy and Buildings*, Elsevier B.V., Vol. 55, p. 459-470.
- Sun, J. et Reddy, T.A. (2006), « Calibration of Building Energy Simulation Programs Using the Analytic Optimization Approach (RP-1051) », *HVAC&R Research*, Vol. 12 No. 1, p. 177-196.
- Taguchi, G. (1986), *Introduction to quality engineering: designing quality into products and processes*, Tokyo : The Organization, p. 191.
- Taillandier, F. (2009), *La notion de risque comme clef du pilotage d'un parc patrimonial immobilier*, Université de Savoie.
- Tarantola, S., Gatelli, D. et Mara, T.A. (2006), « Random balance designs for the estimation of first order global sensitivity indices », *Reliability Engineering & System Safety*, Vol. 91 No. 6, p. 717-727.
- Tian, W. (2013), « A review of sensitivity analysis methods in building energy analysis », *Renewable and Sustainable Energy Reviews*, Elsevier, Vol. 20, p. 411-419.
- Tissot, J.-Y. et Prieur, C. (2012), « Bias correction for the estimation of sensitivity indices based on random balance designs », *Reliability Engineering & System Safety*, Vol. 107, p. 205-213.
- Wang, G.G. (2003), « Adaptive Response Surface Method Using Inherited Latin Hypercube Design Points », *Journal of Mechanical Design*, Vol. 125 No. 2, p. 210.
- Wetter, M. et Polak, E. (2005), « Building design optimization using a convergent pattern search algorithm with adaptive precision simulations », *Energy and Buildings*, Vol. 37 No. 6, p. 603-612.
- Weyl, H. (1938), « Mean Motion », *American Journal of Mathematics*, Vol. 60 No. 4, p. 889.
- Wiener, N. (1938), « The homogeneous chaos », *American Journal of Mathematics*, Vol. 60 No. 4, p. 897-936.

- Wright, J., Wang, M., Brownlee, A. et Buswell, R. (2012), « Variable convergence in evolutionary optimization and its relationship to sensitivity analysis », *Building Simulation and Optimization Conference*, Loughborough, p. 102-109.
- Xiu, D. et Karniadakis, G.E. (2002), « The Wiener--Askey polynomial chaos for stochastic differential equations », *SIAM Journal on Scientific Computing*, Vol. 24 No. 2, p. 619-644.

RESUME

Le secteur du bâtiment est en évolution et se doit de répondre aux problématiques environnementales actuelles aussi bien dans la construction neuve que dans la réhabilitation. La simulation thermique offre le moyen de répondre à cette problématique, mais se limite à une estimation des performances spécifiques aux paramètres qui ont été définis. Ce travail de thèse propose une méthodologie basée sur les puissances de calcul actuelles pour évaluer le comportement du bâtiment sur différentes plages de variation des entrées. Ces variations permettront, en fonction des propriétés qui leur sont attribuées, de rechercher des solutions de réhabilitation ou d'évaluer l'incertitude sur les sorties du modèle thermique.

Afin de réaliser cela, des méthodes d'analyse de sensibilité performantes sont utilisées à travers un outil d'analyse spécifiquement développé pour le logiciel de simulation EnergyPlus. Cet outil permet une évaluation quasi automatique du modèle énergétique de bâtiment via : des techniques d'échantillonnage éprouvées LHS et LP-Tau, des techniques d'analyse performante pour un nombre de simulations réduit RBD-FAST et RBD-SOBOL, une gestion avancée des entrées et sorties permettant une évaluation détaillée du bâtiment, quelle que soit sa géométrie. Les techniques d'analyse de sensibilité sont employées pour proposer une vision précise des principales relations présentes au sein du modèle.

Pour répondre à la contrainte du coût de calcul, la recherche de solution est réalisée à travers un méta-modèle issu d'une décomposition en polynôme de chaos. Une interface graphique utilise cette approximation du modèle complexe du bâtiment pour proposer de façon instantanée une vision claire du comportement de chaque entrée sur les sorties, ainsi que leurs tendances et les plages de variation idéales afin de choisir des solutions. Il est possible de coupler la recherche de solution à l'analyse d'incertitude afin de proposer des solutions robustes et identifier les principales raisons d'un écart entre simulation et réalité.

La méthodologie des travaux de thèse encourage à l'optimisation de la compréhension du modèle plus qu'à la recherche d'une solution particulière. L'objectif de la démarche est de fournir des outils d'analyse afin d'aider l'expert dans la recherche de solution et ainsi de mettre en place des représentations graphiques qui facilitent la compréhension du système bâtiment, pour un gain en transparence et l'intégration de nombreuses contraintes de projet.

Mots Clés : Simulation Thermique Dynamique ; Réhabilitation ; Recherche de performance ; Analyse de sensibilité et d'incertitude ; Méta-modèle

ABSTRACT

The building sector is evolving and has to meet the current environmental issues in both new construction and rehabilitation. The thermal simulation provides the mean to address this problem, but the estimated performance is limited to a specific set of parameters that have been defined. This thesis proposes a methodology based on the increased computing power to assess the building behaviour on different ranges of the inputs variation. Depending on the properties allocated to them, the changes of the variations will allow to seek solutions to restore or to assess the uncertainty in the outputs of the thermal model.

To achieve this, effective sensitivity analysis methods are used across an analysis tool specifically developed for the simulation software EnergyPlus. This tool allows an almost automatic evaluation of the energy model building thanks to the proven sampling techniques such as LHS and LP-Tau; the efficient analysis techniques like RBD-FAST and RBD-SOBOL in order to reduce the number of simulations; an advanced management of the inputs and the outputs for a detailed assessment of the building regardless of its geometry. The sensitivity analysis techniques are used to provide an accurate picture of these key relationships within the model.

To meet the constraints of the computational cost, the search for a solution is achieved through a metamodel from a Polynomial Chaos Expansion PCE. A GUI uses the approximation of the complex model of the building to provide instantly a clear vision of the behaviour of each entry on the outputs, their trends and their ideal ranges of variation to choose solutions. It is possible to couple the search for a solution to the uncertainty analysis to provide robust solutions and identify the main reason for the gap between simulation and reality.

The methodology of the PhD work promotes the optimization of the model understanding more than looking for a specific solution. The purpose of the approach is to provide analytical tools to assist the expert in the search for a solution and so, develop graphical representations that facilitate the understanding of the building system for a gain in transparency and integration of many design constraints.

Keywords: Dynamic Thermal Simulation ; Rehabilitation ; Design performance search ; Sensitivity and uncertainty Analyses ; Metamodel